

CCNP Self-Study
CCNP BSCI Exam
Certification Guide
Third Edition

Clare Gough, CCIE No. 2893

Cisco Press

800 East 96th Street
Indianapolis, IN 46290 USA

CCNP Self-Study CCNP BSCI Exam Certification Guide, Third Edition

Clare Gough

Copyright© 2004 Cisco Press

Published by:

Cisco Press

800 East 96th Street

Indianapolis, IN 46290 USA

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without written permission from the publisher, except for the inclusion of brief quotations in a review.

Printed in the United States of America 1 2 3 4 5 6 7 8 9 0

Library of Congress Cataloging-in-Publication Number: 2002117102

ISBN: 1-58720-085-6

First Printing *December 2003*

Warning and Disclaimer

This book is designed to provide information about BSCI certification. Every effort has been made to make this book as complete and as accurate as possible, but no warranty or fitness is implied.

The information is provided on an “as is” basis. The author, Cisco Press, and Cisco Systems, Inc., shall have neither liability nor responsibility to any person or entity with respect to any loss or damages arising from the information contained in this book or from the use of the discs or programs that may accompany it.

The opinions expressed in this book belong to the author and are not necessarily those of Cisco Systems, Inc.

Feedback Information

At Cisco Press, our goal is to create in-depth technical books of the highest quality and value. Each book is crafted with care and precision, undergoing rigorous development that involves the unique expertise of members from the professional technical community.

Readers’ feedback is a natural continuation of this process. If you have any comments regarding how we could improve the quality of this book, or otherwise alter it to better suit your needs, you can contact us through e-mail at feedback@ciscopress.com. Please make sure to include the book title and ISBN in your message.

We greatly appreciate your assistance.

Corporate and Government Sales

Cisco Press offers excellent discounts on this book when ordered in quantity for bulk purchases or special sales. For more information, please contact: **U.S. Corporate and Government Sales** 1-800-382-3419 corpsales@pearsontechgroup.com

For sales outside of the U.S. please contact: **International Sales** 1-317-581-3793 international@pearsontechgroup.com

Trademark Acknowledgments

All terms mentioned in this book that are known to be trademarks or service marks have been appropriately capitalized. Cisco Press or Cisco Systems, Inc., cannot attest to the accuracy of this information. Use of a term in this book should not be regarded as affecting the validity of any trademark or service mark.

Publisher: John Wait

Editor-in-Chief: John Kane

Executive Editor: Brett Bartow

Cisco Representative: Anthony Wolfenden

Cisco Press Program Manager: Sonia Torres Chavez

Manager, Marketing Communications,
Cisco Systems: Scott Miller

Cisco Marketing Program Manager: Edie Quiroz

Production Manager: Patrick Kanouse

Acquisitions Editor: Michelle Grandin

Development Editor: Dayna Isley

Project Editor: Argosy Publishing

Copy Editor: Ginny Kaczmarek

Technical Editors: Tim Sammut, Gert Schauwers, Brent Stewart

Team Coordinator: Tammi Barnett

Book and Cover Designer: Louisa Adair

Composition: Argosy Publishing

Indexer: Larry Sweazy



Corporate Headquarters

Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 526-4100

European Headquarters

Cisco Systems International BV
Haarlerbergpark
Haarlerbergweg 13-19
1101 CH Amsterdam
The Netherlands
www-europe.cisco.com
Tel: 31 0 20 357 1000
Fax: 31 0 20 357 1100

Americas Headquarters

Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-7660
Fax: 408 527-0883

Asia Pacific Headquarters

Cisco Systems, Inc.
Capital Tower
168 Robinson Road
#22-01 to #29-01
Singapore 068912
www.cisco.com
Tel: +65 6317 7777
Fax: +65 6317 7799

Cisco Systems has more than 200 offices in the following countries and regions. Addresses, phone numbers, and fax numbers are listed on the

Cisco.com Web site at www.cisco.com/go/offices.

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China PRC • Colombia • Costa Rica • Croatia • Czech Republic
Denmark • Dubai, UAE • Finland • France • Germany • Greece • Hong Kong SAR • Hungary • India • Indonesia • Ireland • Israel • Italy
Japan • Korea • Luxembourg • Malaysia • Mexico • The Netherlands • New Zealand • Norway • Peru • Philippines • Poland • Portugal
Puerto Rico • Romania • Russia • Saudi Arabia • Scotland • Singapore • Slovakia • Slovenia • South Africa • Spain • Sweden
Switzerland • Taiwan • Thailand • Turkey • Ukraine • United Kingdom • United States • Venezuela • Vietnam • Zimbabwe

Copyright © 2003 Cisco Systems, Inc. All rights reserved. CCIP, CCSP, the Cisco Arrow logo, the Cisco *Powered* Network mark, the Cisco Systems Verified logo, Cisco Unity, Follow Me Browsing, FormShare, iQ Net Readiness Scorecard, Networking Academy, and ScriptShare are trademarks of Cisco Systems, Inc.; Changing the Way We Work, Live, Play, and Learn, The Fastest Way to Increase Your Internet Quotient, and iQuick Study are service marks of Cisco Systems, Inc.; and Aironet, ASIST, BPX, Catalyst, CCDA, CCDP, CCIE, CCNA, CCNP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, the Cisco IOS logo, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Empowering the Internet Generation, Enterprise/Solver, EtherChannel, EtherSwitch, Fast Step, GigaStack, Internet Quotient, IOS, IPTV, iQ Expertise, the iQ logo, LightStream, MGX, MICA, the Networkers logo, Network Registrar, *Packet*, PIX, Post-Routing, Pre-Routing, RateMUX, Registrar, SlideCast, SMARTnet, StrataView Plus, Stratum, SwitchProbe, TeleRouter, TransPath, and VCO are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the U.S. and certain other countries.

All other trademarks mentioned in this document or Web site are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (0303R)

Printed in the USA

About the Author

Clare Gough, CCIE No. 2893, has worked as a network engineer for 17 years, during which time she has managed and designed large-scale networks. In 1997 she was one of the first women to gain her CCIE.

In addition to being a network design consultant, she is an acknowledged authority on the development and delivery of technical training. She is the author of several Cisco Press books, now used throughout the world. At Digital Equipment Corporation in Atlanta, she was part of the development team responsible for the production of all training materials. She also created PATHworks System Management courses. As the third employee hired by Protocol Interface, which was the first Cisco Training Partner, Clare eventually became the company's first training manager and saw it grow to become a part of Global Knowledge.

Born in England, Clare holds two master's degrees and moved to the United States in 1991, where she now lives with her husband and young son.

About the Technical Reviewers

Tim Sammut, CCIE No. 6642, is a senior network consultant for Northrop Grumman Information Technology. Tim has served in key project roles involving technologies from LAN switching to security to SNA integration, and has helped many organizations, ranging from 100 to 130,000 users, make the most of their network investment. Tim also holds the CISSP, CCIE Security, and CCIE Communication and Services certifications.

Gert Schauwer, CCIE No. 6924, has CCIE certifications in Security, Routing and Switching, and Communications and Service. He has more than five years of experience in internetworking. He is currently working for the CCIE team at Cisco in Brussels, Belgium, as a CCIE content engineer. He has an engineering degree in electronics.

Brent Stewart is a certified Cisco Systems instructor for Global Knowledge. With Global Knowledge, he has participated in the development of ICND, BSCI, BCMSN, BCRAN, and CIT for Internet-based delivery and served as SME (Subject Matter Expert) for the CD-based ICND and CIT titles. He helped author the BSCI 2.0 update and served as the lab development engineer. Before working for Global Knowledge, Brent owned an ISP and worked as an IT consultant. Brent holds the CCNP, CCDP, and MCSE certifications. Brent lives in Hickory, North Carolina, with his wife, Karen, and beautiful but mischievous children, Benjamin, Kaitlyn, and Madelyn.

Dedications

It has been a tough year for many people, where family and friends have proved to be the foundation to everything. To all those people, I thank you, but in particular to my husband David, whose advice and care I value above all else, and to my young scamp of a son, who provided no practical help in the writing of this book, but much joy and endless entertainment.

Acknowledgments

While watching the Oscars, I used to wonder why people cannot simply say “Thanks a bunch” and let us find out who is the next lucky recipient, but after writing books, I realize how much work goes into such a production.

All books are the product of a team, of which I am just one player. I have been blessed with a dedicated and professional team. As with the other books that I have written for Cisco Press, I have come to think of the team members as friends as well as colleagues, whose humor and consideration have made this such an enjoyable project. In particular, I would like to thank John Kane, the editor-in-chief, who I continue to enjoy working with and whose friendship I value. Michelle Grandin, the acquisitions editor, provided creative solutions and endless support, understanding, and flexibility. She encouraged me during those dark moments and provided light at the end of the tunnel.

Dayna Isley, the development editor, kept me focused and taught me so much about the art of book writing, dotting my i’s and crossing my t’s. Patrick Kanouse, the production manager, turned the raw manuscript into the book you see before you and, where necessary, even painstakingly deciphering my scrawl.

The technical editors on this project have been quite brilliant; to them lies the credit for the completeness and technical strength of the book. Each of the editors lent his particular skills to this project. Dave Hucaby offered amazing attention to detail and suggestions of areas requiring clarification. Tim Sammut’s eagle eyes and technical expertise have kept me honest. Gert Schauwers added great value as the resident Cisco Systems expert.

I would particularly like to thank Brent Stewart, whose insights about where to add more detail or remove errors were invaluable. He provided many of the configuration examples in the book and, most importantly, laughter and friendship.

There are many people behind the scenes who go unacknowledged, but whose work is crucial to the success of such a project. Thank you to all of these unsung heroes.

Contents at a Glance

Foreword xxv

Introduction: All About the CCNP, CCDP, and CCIP Certifications xxvi

Part I IP Routing Fundamentals 3

Chapter 1 IP Routing Principles 5

Chapter 2 IP Addressing 37

Chapter 3 Designing IP Networks 85

Chapter 4 IP Distance Vector Routing Principles 133

Chapter 5 IP Link-State Routing Principles 155

Part II OSPF 177

Chapter 6 Using OSPF in a Single Area 179

Chapter 7 Configuring OSPF in a Single Area 219

Chapter 8 Using OSPF Across Multiple Areas 263

Chapter 9 Configuring OSPF Across Multiple Areas 291

Part III IS-IS 333

Chapter 10 Fundamentals of the Integrated IS-IS Protocol 335

Chapter 11 Integrated IS-IS Protocol Operation 375

Chapter 12 Configuring Integrated IS-IS 393

Part IV EIGRP 439

Chapter 13 Using EIGRP in Enterprise Networks 441

Chapter 14 Configuring EIGRP in Enterprise Networks 475

Part V BGP 507

Chapter 15 Connecting to Other Autonomous Systems—The Basics of BGP 509

Chapter 16 Implementing and Tuning BGP for Use in Large Networks 557

Part VI Redistribution and Policy-Based Routing 611

Chapter 17 Implementing Redistribution and Controlling Routing Updates 613

Chapter 18 Controlling Network Traffic with Route Maps and Policy-Based Routing 665

Part VII Scenarios 697

Chapter 19 Scenarios for Exam Preparation 699

Part VIII Appendixes 765

Appendix A Answers to Chapter “Do I Know This Already?” Quizzes and Q&A
Sections 767

Appendix B Suggested Reading 845

Glossary 851

Index 866

Contents

Foreword xxv

Introduction: All About the CCNP, CCDP, and CCIP Certifications xxvi

Part I IP Routing Fundamentals 3

Chapter 1 IP Routing Principles 5

“Do I Know This Already?” Quiz 5

Foundation Topics 9

Understanding Routing Protocol Fundamentals 9

The Definition of a Routing Protocol 9

How the Routing Protocol Works 9

Routing and Routed 10

Types of Routing Protocols 11

Classful Routing 11

Classless Routing 12

The ip classless Command 12

The Routing Table 13

The Routing Table Fields 14

The Network Field 14

The Outgoing Interface Field 15

The Metric Field 15

The Next Hop Field 16

How to Keep the Routing Table Current and Correct 16

The show ip route Command 17

Other Methods of Entering Routes into the Routing Table 19

Static Routes 20

Default Static Routes 21

Floating Static Routes 22

On Demand Routing 23

Routing Versus Switching 25

The Routing Function 25

The Switching Function 26

The Routing/Switching Relationship in a Cisco Router 26

Foundation Summary 28

Q&A 30

Scenarios 32

Scenario 1-1 32

Scenario Answers 34

Scenario 1-1 Answers 34

Chapter 2 IP Addressing 37

“Do I Know This Already?” Quiz 37

Foundation Topics 41

IP Subnetting	41
<i>The Need for Layer 3 Addressing</i>	41
<i>Network Structures and Data Flow</i>	41
<i>The Network and How It Is Addressed</i>	41
<i>Layer 3 Network Characteristics</i>	42
<i>An Analogy for Understanding the Concept of a Network</i>	43
<i>Layer 3 to Layer 2 Conversion</i>	43
<i>An IP Address</i>	44
<i>The Internet Authoritative Bodies</i>	46
<i>An Example of Bit Allocation in a Network Address</i>	47
<i>The Subnet Mask</i>	47
<i>The Logical AND</i>	47
<i>Familiar Rules in IP Subnetting</i>	48
<i>The New Subnet Rules</i>	49
Prefix Routing/CIDR	50
<i>A Definition of Prefix Routing/CIDR</i>	51
<i>Problems with IP Addressing and the Internet</i>	51
<i>CIDR as a Solution</i>	52
<i>An Example of the Use of CIDR</i>	53
<i>Advantages of Prefix Routing/CIDR</i>	54
Variable-Length Subnet Masks	56
<i>An Example of VLSM</i>	56
<i>Rules for VLSM</i>	57
<i>Case Study: Addressing the Network</i>	57
<i>Allocating VLSM Addresses</i>	60
<i>Optimizing the IP Address Space</i>	61
<i>Assigning IP VLSM Subnets for WAN Connections</i>	62
Summarization	64
<i>The Advantages of Summarization</i>	67
<i>Reducing the Size of the Routing Table</i>	67
<i>Hiding Network Changes</i>	67
<i>Other Solutions to Address Exhaustion</i>	68
<i>Configuring Summarization</i>	68
<i>Automatic Summarization</i>	68
<i>Manual Summarization</i>	69
Foundation Summary	72
Q&A	75
Scenarios	77
Scenario 2-1	77
Scenario 2-2	77
Scenario Answers	79
Scenario 2-1 Answers	79
Scenario 2-2 Answers	82

Chapter 3	Designing IP Networks	85
	“Do I Know This Already?” Quiz	85
	Foundation Topics	90
	Criteria in Designing IP Networks	90
	<i>The Cisco Hierarchical Design</i>	90
	<i>The Functions of Each Layer</i>	91
	<i>General Design Rules for Each Layer</i>	92
	IP Access Lists	93
	Security Using Access Lists	94
	Controlling Terminal Access	94
	Traffic Control Through Routing Updates	94
	Prioritization	95
	Reducing Network Traffic: Alternatives to Access Lists	96
	Internet Example	97
	Intranet Example	98
	Keys Points to Remember When Designing an IP Network	99
	Designing IP Addresses for an Existing Network	101
	Private Addresses on the Internet	103
	Connecting to the Outside World with NAT	105
	The Main Features of NAT	107
	The Main Functions of NAT	107
	Understanding IPv6	108
	IPv6 Address Format	110
	IPv6 Unicast Addresses	111
	IPv6 Multicast Addresses	112
	Address Aggregation	112
	Autoconfiguration	113
	Renumbering	114
	Simple and Efficient Header	114
	IPv6 Extension Header	114
	Security	115
	Mobility	115
	IPv4 to IPv6 Transitions	116
	IPv6 Routing Protocols	117
	Foundation Summary	118
	Q&A	122
	Scenarios	123
	Scenario 3-1	123
	Scenario Answers	126
	Scenario 3-1 Answers	126
Chapter 4	IP Distance Vector Routing Principles	133
	“Do I Know This Already?” Quiz	133
	Foundation Topics	138

	Distance Vector Routing Protocols	138
	<i>Operation of Distance Vector Routing Protocols</i>	138
	<i>The Distance Vector Routing Metrics</i>	139
	RIP Version 1 and RIP Version 2	140
	RIPv1	140
	RIPv2	142
	IGRP and EIGRP	143
	IGRP	143
	EIGRP	145
	Choosing Between Routing Protocols Using Administrative Distance	146
	Convergence	148
	RIPv1 and RIPv2 Convergence	149
	IGRP Convergence	149
	EIGRP Convergence	150
	Interior and Exterior Gateway Protocols	150
	Foundation Summary	151
	Q&A	153
Chapter 5	IP Link-State Routing Principles	155
	“Do I Know This Already?” Quiz	155
	Foundation Topics	160
	Link-State Routing Protocol Overview	160
	<i>The Meaning of Link State</i>	160
	<i>Learning About the Network</i>	160
	<i>Updating Local Network Tables</i>	161
	<i>Path Selection</i>	161
	OSPF	161
	IS-IS	163
	BGP-4	163
	Convergence	164
	<i>OSPF Convergence</i>	165
	<i>IS-IS Convergence</i>	165
	<i>BGP Convergence</i>	165
	Foundation Summary	166
	Q&A	170
	Scenarios	171
	Scenario 5-1	171
	Scenario Answers	173
	Scenario 5-1 Answers	173
Part II	OSPF	177
Chapter 6	Using OSPF in a Single Area	179
	“Do I Know This Already?” Quiz	179

Foundation Topics 184

Understanding OSPF Fundamentals 184

OSPF Terminology 184

OSPF Features 186

*OSPF Neighbors 187**Adjacent OSPF Neighbors 187**Advantages of Having Neighbors 187**The DR 188**BDRs 188**Electing the DRs and BDRs 189**The Hello Packet 191*

OSPF Operation in a Single Area 193

*Creating and Maintaining the OSPF Routing Table 193**How OSPF Builds the Routing Table on a New Router 194**Finding Neighbors with the Exchange Process 195**Discovering Routes 196**The Topology Database 197**Maintaining the Topological Database and the Routing Table 199**Learning a New Route 200**Choosing the Shortest Path First and Building the Routing Table 201**The Metric 201**Information Needed in the Routing Table 201*

OSPF Network Topologies 202

*Broadcast Multiaccess Network 202**Point-to-Point Network 203**Point-to-Multipoint Network 203**Nonbroadcast Multiaccess Network 203**Virtual Links 204**OSPF Across NBMA Networks 205**Choosing a Topology 206**Subinterfaces 206***Foundation Summary 208****Q&A 214****Scenarios 216**

Scenario 6-1 216

Scenario Answers 217

Scenario 6-1 Answers 217

Chapter 7 Configuring OSPF in a Single Area 219

“Do I Know This Already?” Quiz 219

Foundation Topics 224

Configuring OSPF in a Single Area 224

*Required Commands for Configuring OSPF on an Internal Router 224**Enabling the OSPF Routing Protocol 225*

	<i>Enabling the OSPF network Command</i>	225
	<i>Options for Configuring OSPF on an Internal Router</i>	228
	<i>The Loopback Interface and the Router ID</i>	228
	<i>Changing the Default Metric Using the cost Command</i>	229
	<i>Determining the Designated Router Using the priority Command</i>	231
	<i>A Working Configuration of OSPF on a Single Router</i>	232
	Configuring OSPF over an NBMA Topology	232
	<i>Configuring OSPF in NBMA Mode</i>	234
	<i>Configuring OSPF in Point-to-Multipoint Mode</i>	235
	<i>Configuring OSPF in Broadcast Mode</i>	236
	<i>Configuring OSPF in Point-to-Point Mode on a Frame Relay Subinterface</i>	236
	Checking the Configuration of OSPF on a Single Router	237
	<i>The show ip ospf Command</i>	238
	<i>The show ip ospf database Command</i>	240
	<i>The show ip ospf interface Command</i>	242
	<i>The show ip ospf neighbor Command</i>	244
	<i>The show ip protocols Command</i>	247
	<i>The show ip route Command</i>	249
	Troubleshooting OSPF in a Single Area	249
	<i>The debug Commands</i>	250
	Foundation Summary	251
	Q&A	253
	Scenarios	255
	Scenario 7-1	255
	Scenario Answers	256
	Scenario 7-1 Answers	256
Chapter 8	Using OSPF Across Multiple Areas	263
	“Do I Know This Already?” Quiz	263
	Foundation Topics	268
	The Purpose of Using OSPF in a Multiple Area Network	268
	<i>Problems with OSPF in a Single Area</i>	268
	<i>How to Determine Area Boundaries</i>	269
	The Features of Multiple Area OSPF	269
	<i>OSPF Within an Area</i>	269
	<i>Router Types</i>	270
	<i>Link-State Advertisements</i>	272
	<i>The Different Types of Areas</i>	272
	The Operation of OSPF Across Multiple Areas	274
	<i>The ABRs and ASBR Propagation of LSAs</i>	274
	<i>OSPF Path Selection Between Areas</i>	274
	<i>Calculating the Cost of a Path to Another Area</i>	276
	Design Considerations in Multiple Area OSPF	277
	<i>Capacity Planning in OSPF</i>	278

	<i>Number of Neighbors per Router</i>	279
	<i>Number of Areas per ABR</i>	280
	<i>Summarization</i>	280
	<i>The Virtual Link</i>	280
	<i>Multiple Area OSPF Over an NBMA Network</i>	281
	Foundation Summary	283
	Q&A	286
	Scenarios	287
	Scenario 8-1	287
	Scenario Answers	289
	Scenario 8-1 Answers	289
Chapter 9	Configuring OSPF Across Multiple Areas	291
	“Do I Know This Already?” Quiz	291
	Foundation Topics	296
	Required Configuration Commands for a Multiarea OSPF Network	296
	<i>Enabling the OSPF Routing Protocol</i>	296
	<i>Enabling the network Command</i>	297
	Optional Configuration Commands for a Multiarea OSPF Network	298
	<i>The area range Command</i>	299
	<i>The summary-address Command</i>	300
	<i>The area area-id stub Command</i>	302
	<i>The area area-id stub no-summary Command</i>	303
	<i>The area default-cost Command</i>	303
	<i>The area virtual-link Command</i>	305
	A Working Configuration of OSPF on a Multiarea Network	308
	Verifying the Configuration of OSPF in a Multiarea Network	309
	<i>The show ip ospf border-routers Command</i>	310
	<i>The show ip route Command</i>	311
	<i>The show ip ospf virtual-links Command</i>	312
	<i>The show ip ospf database Command</i>	313
	Troubleshooting a Multiarea OSPF Network	315
	<i>The log-adjacency-changes Command</i>	315
	<i>Useful debug Commands</i>	316
	<i>The debug ip packet Command</i>	317
	<i>The debug ip ospf events Command</i>	317
	<i>Common Problems with Adjacencies</i>	317
	Foundation Summary	318
	Q&A	320
	Scenarios	322
	Scenario 9-1	322
	Scenario 9-2	323
	Scenario 9-3	326

Scenario Answers 327

Scenario 9-1 Answers 327

Scenario 9-2 Answers 329

Scenario 9-3 Answers 331

Part III IS-IS 333**Chapter 10 Fundamentals of the Integrated IS-IS Protocol 335**

“Do I Know This Already?” Quiz 335

Foundation Topics 341

Introduction to Integrated IS-IS 341

IS-IS Terminology 341

OSPF and IS-IS Compared 346

*Similarities Between Integrated IS-IS and OSPF 346**Differences Between OSPF and Integrated IS-IS 348*

ISO Addressing for Integrated IS-IS 350

*NETs and NSAP 352**Rules of ISO Addressing 353**Example of a NET Address 353*

Integrated IS-IS Hierarchical Structure 354

*The Level 1 Router 354**The Level 2 Router 354**The Level 1-2 Router 355*

Basic Principles of Area Routing 355

Integrated IS-IS Networks and Interfaces 357

*Establishing Adjacencies on a Point-to-Point Link 358**Establishing Adjacencies on a Broadcast Link 358**Establishing Adjacencies on an NBMA Link 359*

Network Layer Protocols Used in Integrated IS-IS 360

*The Format of the Hello Packet 361**The Format of the LSP 363**The Format of the SNP 365**TLVs 365***Foundation Summary 368****Q&A 372****Chapter 11 Integrated IS-IS Protocol Operation 375**

“Do I Know This Already?” Quiz 375

Foundation Topics 379

Integrated IS-IS Operation 379

*The Update Process 379**Sending and Receiving an LSP 380**Determining Whether the LSP in the Database Is Valid 382**The Decision Process 382**The Forwarding and Receive Processes 385*

Integrated IS-IS Design Considerations	385
<i>Area Design of Integrated IS-IS Routers</i>	385
<i>Route Summarization</i>	387
<i>Integrated IS-IS NBMA Modeling Solutions in a Switched WAN Network</i>	387

Foundation Summary 388

Q&A 390

Chapter 12 Configuring Integrated IS-IS 393

“Do I Know This Already?” Quiz	393

Foundation Topics 397

Basic Configuration of Integrated IS-IS	397
Optional Integrated IS-IS Commands	399
<i>Changing the Router from Level 1-2</i>	399
<i>Configuring Summarization</i>	401
<i>Configuring NBMA</i>	403
<i>The Broadcast Configuration over NBMA</i>	404
<i>The Point-to-Point Configuration over NBMA</i>	406
Verifying the Integrated IS-IS Operation	407
<i>The show clns neighbors Command</i>	409
<i>The show clns interface Command</i>	412
<i>The show isis database Command</i>	414
<i>The show isis database detail Command</i>	416
Troubleshooting the Integrated IS-IS Operation	419
<i>The show isis spf-log Command</i>	419
<i>The debug Commands</i>	421

Foundation Summary 423

Q&A 425

Scenarios 426

Scenario 12-1	426
Scenario 12-2	427

Scenario Answers 431

Scenario 12-1 Answers	431
Scenario 12-2 Answers	435

Part IV EIGRP 439

Chapter 13 Using EIGRP in Enterprise Networks 441

“Do I Know This Already?” Quiz	441

Foundation Topics 445

Overview of EIGRP in an Enterprise Network	445
<i>EIGRP Terminology</i>	445
<i>Understanding EIGRP Features and Advantages</i>	447
<i>DUAL</i>	448
<i>Loop-Free Networks</i>	448
<i>Incremental Updates</i>	448

<i>Multicast Addressing for Updates</i>	448
<i>Advanced Distance Vector Protocol</i>	449
<i>Loop-Free Routing Tables</i>	449
<i>Support for Different Topologies</i>	449
<i>Rapid Convergence</i>	449
<i>Reduced Bandwidth Use</i>	449
<i>Protocol Independence at Layer 3</i>	449
<i>Compatibility with IGRP</i>	450
<i>Easy Configuration</i>	450
<i>Use of a Composite Metric</i>	450
<i>Unequal-Cost Load Balancing</i>	450
<i>Components of EIGRP</i>	450
<i>Protocol-Dependent Modules</i>	451
<i>RTP</i>	451
<i>Neighbor Discovery and Recovery</i>	451
<i>DUAL</i>	451
The Operation of EIGRP	451
<i>Creating the Neighbor Table</i>	452
<i>The Contents of the Neighbor Table</i>	452
<i>Becoming a Neighbor</i>	453
<i>Creating the Topology Table</i>	454
<i>Maintaining the Topology Table</i>	455
<i>Adding a Network to the Topology Table</i>	456
<i>Removing a Path or Router from the Topology Table</i>	458
<i>Finding an Alternative Path to a Remote Network</i>	458
<i>Creating the Routing Table</i>	460
<i>EIGRP Metrics</i>	460
<i>The Topology Table and the DUAL Finite-State Machine</i>	462
<i>Updating the Routing Table in Passive Mode with DUAL</i>	462
<i>Updating the Routing Table in Active Mode with DUAL</i>	464
<i>Choosing a Successor</i>	465
EIGRP Network Design	466
<i>Solutions to EIGRP Scaling Issues</i>	467
<i>EIGRP Design Issues</i>	467
Foundation Summary	469
Q&A	473
Chapter 14 Configuring EIGRP in Enterprise Networks	475
“Do I Know This Already?” Quiz	475
Foundation Topics	479
Configuring EIGRP	479
<i>Required Commands for Configuring EIGRP</i>	479
<i>Optional Commands for Configuring EIGRP</i>	481
<i>Summarization with EIGRP</i>	482
<i>Stub Routers</i>	483

<i>Load Balancing in EIGRP</i>	485
<i>Tuning the EIGRP Process</i>	487
<i>Optional EIGRP Commands over WANs</i>	489
<i>EIGRP Defaults in Bandwidth Utilization</i>	489
<i>Configuring the Pure Point-to-Point Network</i>	491
Verifying the EIGRP Operation	491
<i>The show ip eigrp neighbors Command</i>	492
<i>The show ip eigrp topology Command</i>	493
<i>The show ip eigrp traffic Command</i>	495
Troubleshooting the EIGRP Operation	496
Foundation Summary	497
Q&A	499
Scenarios	501
Scenario 14-1	501
Scenario 14-2	502
Scenario Answers	504
Scenario 14-1 Answers	504
Scenario 14-2 Answers	505
 Part V BGP	 507
Chapter 15 Connecting to Other Autonomous Systems—The Basics of BGP	509
“Do I Know This Already?” Quiz	509
Foundation Topics	513
Introduction to BGP and Communicating with Other Autonomous Systems	513
<i>Characteristics of BGP</i>	513
<i>BGP Terminology</i>	515
<i>When to Use BGP</i>	517
<i>When Not to Use BGP</i>	518
Overview of the BGP Operation	518
<i>CIDR and Route Aggregation</i>	519
<i>BGP and Policy-Based Routing</i>	521
<i>Rules of Policy-Based Routing</i>	523
<i>BGP Attributes</i>	523
<i>Route Selection Process</i>	529
Basic Configuration Commands to Connect to Another Autonomous System	530
<i>Configuring Required BGP Commands</i>	530
<i>Starting the Routing Process</i>	530
<i>Identifying Neighbors and Defining Peer Groups</i>	530
<i>Configuring Optional BGP Commands</i>	532
<i>Defining the Networks to Be Advertised</i>	532
<i>Forcing the Next-Hop Address</i>	533
<i>Aggregating Routes</i>	533
Managing and Verifying the BGP Configuration	535

Alternative Methods of Connecting to Another Autonomous System	536
Foundation Summary	538
Q&A	545
Scenarios	547
Scenario 15-1	547
Scenario 15-2	548
Scenario 15-3	549
Scenario Answers	552
Scenario 15-1 Answers	552
Scenario 15-2 Answers	553
Scenario 15-3 Answers	554
Chapter 16 Implementing and Tuning BGP for Use in Large Networks	557
“Do I Know This Already?” Quiz	557
Foundation Topics	562
Building a Network Using Internal BGP	562
Understanding iBGP Network Requirements	562
<i>Physical versus Logical Connections</i>	562
<i>Synchronization</i>	564
<i>A Fully Meshed Network</i>	566
<i>Network Resources Required in Fully Meshed Networks</i>	566
<i>Administrative Overhead in Fully Meshed Networks</i>	566
Designing and Configuring a Scalable iBGP Network	566
<i>Route Reflectors</i>	567
<i>How Route Reflectors Operate</i>	570
<i>Configuring Route Reflectors</i>	571
<i>Route Refresh</i>	573
<i>Peer Groups</i>	574
Verifying the iBGP Configuration	575
Controlling BGP Traffic	578
<i>How Prefix Lists Work</i>	579
<i>How to Configure a BGP Prefix List</i>	580
<i>Verifying the Prefix List Configuration</i>	581
Connecting to the Internet with BGP	583
<i>Redundant Connections to the Internet—Multihoming</i>	583
<i>Receiving Routing Information from the Internet</i>	584
Determining the BGP Path by Tuning the Attributes	586
<i>Using the Local Preference and Weight Attributes</i>	586
<i>Verifying the Configuration of Attributes</i>	590
Redistribution Between IGP and BGP	593
<i>Advertising Routes from IGP into BGP</i>	593
<i>Advertising Routes from BGP into an IGP</i>	593
Foundation Summary	596

Q&A 600**Scenarios 602**

Scenario 16-1 602

Scenario 16-2 603

Scenario 16-3 604

Scenario Answers 606

Scenario 16-1 Answers 606

Scenario 16-2 Answers 607

Scenario 16-3 Answers 608

Part VI Redistribution and Policy-Based Routing 611**Chapter 17 Implementing Redistribution and Controlling Routing Updates 613**

“Do I Know This Already?” Quiz 613

Foundation Topics 618

Understanding the Fundamentals of Redistribution 618

Understanding the Routing Decisions That Affect Redistribution 621

*Routing Metrics and Redistribution 622**Path Selection Between Routing Protocols 623**Avoiding Routing Loops When Redistributing 624**Avoiding Suboptimal Routing Decisions When Redistributing 626**Avoiding Problems with Network Convergence When Redistributing 627*

Controlling Routing Updates During Redistribution 628

*Passive Interfaces 628**Static Routes 628**Default Routes 629**The null Interface 630**Distribute Lists 631**Route Maps 631*

Configuring Redistribution 632

*Redistribution Configuration Syntax 633**Configuring the Default Metric 635**Configuring the Default Metric for OSPF, IS-IS, RIP, EGP, or BGP 635**Configuring the Default Metric for EIGRP or IGRP 636**Configuring the Administrative Distance 637**Configuration Commands to Control Routing Updates in Redistribution 639**Configuring the Passive Interface 639**Configuring Static Routes 639**Configuring Default Routes 641**Redistribution Examples 642**Example 1: Route Redistribution Without Redundant Paths 642**Example 2: Route Redistribution with Redundant Paths 644**Example 3: A Default Network in a Redistributed Environment 647*

Controlling Routing Updates with Filtering 648

Verifying, Maintaining, and Troubleshooting the Implementation of Redistribution and Filtering 650

The trace Command 651

The Extended ping Command 651

Using trace and Extended ping 651

Foundation Summary 652

Q&A 656

Scenario 658

Scenario 17-1 658

Scenario Answers 660

Scenario 17-1 Answers 660

Chapter 18 Controlling Network Traffic with Route Maps and Policy-Based Routing 665

“Do I Know This Already?” Quiz 665

Foundation Topics 671

Understanding Route Maps 671

Understanding Policy-Based Routing 673

Benefits of Policy-Based Routing 674

Disadvantages of Policy-Based Routing 675

The Operation of Route Maps and Policy-Based Routing 675

Configuring Route Maps for Policy-Based Routing 677

The match Commands for Policy-Based Routing with Route Maps 678

The set Commands for Policy-Based Routing with Route Maps 678

Configuring Fast Switching with Policy-Based Routing 680

Configuring Route Maps for Redistribution 681

The match Commands for Redistribution with Route Maps 681

The set Commands for Redistributing with Route Maps 682

Monitoring the Configuration of Route Maps, Policy-Based Routing, and Redistribution 684

Foundation Summary 685

Q&A 691

Scenarios 692

Scenario 18-1 692

Scenario Answers 694

Scenario 18-1 Answers 694

Part VII Scenarios 697

Chapter 19 Scenarios for Exam Preparation 699

Further Study for Final Exam Preparation 699

How to Best Use This Chapter 700

Scenarios 701

Scenario 19-1 701

Scenario 19-2 702

Scenario 19-3 703

Scenario Answers 705

Scenario 19-1 Answers 705

Scenario 19-2 Answers 707

Scenario 19-3 Answers 708

Three-Part Scenarios 712

Scenario 19-4 712

*Scenario 19-4, Part A—Planning 712**Solutions to Scenario 19-4, Part A—Planning 718**Scenario 19-4, Part B—Configuration 723**Solutions to Scenario 19-4, Part B—Configuration 723**Scenario 19-4, Part C—Verification and Questions 724**Solutions to Scenario 19-4, Part C—Verification and Questions 728*

Scenario 19-5 731

*Scenario 19-5, Part A—Planning 732**Solutions to Scenario 19-5, Part A—Planning 732**Scenario 19-5, Part B—Configuration 735**Solutions to Scenario 19-5, Part B—Configuration 735**Scenario 19-5, Part C—Verification and Questions 738**Solutions to Scenario 19-5, Part C—Verification and Questions 741*

Scenario 19-6 746

*Scenario 19-6, Part A—Planning 746**Solutions to Scenario 19-6, Part A—Planning 752**Scenario 19-6, Part B—Configuration 753**Solutions to Scenario 19-6, Part B—Configuration 753**Scenario 19-6, Part C—Verification and Questions 756**Solutions to Scenario 19-6, Part C—Verification and Questions 760***Part VIII Appendixes 765**Appendix A Answers to Chapter “Do I Know This Already?” Quizzes and Q&A
Sections 767

Appendix B Suggested Reading 845

Glossary 851

Index 866

Foreword

CCNP BSCI Exam Certification Guide, Third Edition, is a complete study tool for the CCNP BSCI exam, enabling you to assess your knowledge, identify areas to concentrate your study, and master key concepts to help you succeed on the exams and in your daily job. The book is filled with features that help you master the skills necessary to design and configure routing protocols based on the principles of IP routing. This book was developed in cooperation with the Cisco Internet Learning Solutions Group. Cisco Press books are the only self-study books authorized by Cisco for CCNP exam preparation.

Cisco and Cisco Press present this material in text-based format to provide another learning vehicle for our customers and the broader user community in general. Although a publication does not duplicate the instructor-led or e-learning environment, we acknowledge that not everyone responds in the same way to the same delivery mechanism. It is our intent that presenting this material via a Cisco Press publication will enhance the transfer of knowledge to a broad audience of networking professionals.

Cisco Press will present study guides on existing and future exams through these Exam Certification Guides to help achieve Cisco Internet Learning Solutions Group's principal objectives: to educate the Cisco community of networking professionals and to enable that community to build and maintain reliable, scalable networks. The Cisco career certifications and classes that support these certifications are directed at meeting these objectives through a disciplined approach to progressive learning. To succeed on the Cisco career certifications exams, as well as in your daily job as a Cisco-certified professional, we recommend a blended learning solution that combines instructor-led, e-learning, and self-study training with hands-on experience. Cisco Systems has created an authorized Cisco Learning Partner program to provide you with the most highly qualified instruction and invaluable hands-on experience in lab and simulation environments. To learn more about Cisco Learning Partner programs available in your area, please go to www.cisco.com/go/authorizedtraining.

The books Cisco Press creates in partnership with Cisco Systems will meet the same standards for content quality demanded of our courses and certifications. It is our intent that you will find this and subsequent Cisco Press certification and training publications of value as you build your networking knowledge base.

Thomas M. Kelly
Vice-President, Internet Learning Solutions Group
Cisco Systems, Inc.
October 2003

Introduction: All About the CCNP, CCDP, and CCIP Certifications

The CCNP, CCDP, and CCIP certifications are becoming increasingly popular. These certifications, which have as their foundation the CCNA certification, form the second rung in the ladder to the coveted CCIE certification. The BSCI exam (642-801) is one of the exams that you must pass to become a CCNP, CCDP, or CCIP. This book will help you prepare for that exam. Professional-level certification opens doors to career opportunities and is a prerequisite for other Cisco certifications, as well. Generally, passing the BSCI exam means that you have mastered the concepts and implementation skills necessary to build a complex IP network of Cisco routers.

NOTE The CCNP, CCDP, and CCIP certifications are often referred to as the *professional-level certifications*. For more information on the differences among the three professional-level certifications and the latest on Cisco exams and certifications, begin at the Cisco Career Certification page (<http://www.cisco.com/go/training>) at the Cisco web site. From there, you can navigate to the BSCI exam.

The BSCI exam is a computer-based exam, with multiple-choice, fill-in-the-blank, and list-in-order questions. You can take the exam at any Prometric (<http://www.2test.com>) or Pearson VUE testing center (<http://www.pearsonvue.com>). Your testing center can tell you the exact length of the exam. (Be aware that when you register for the exam, you might be told to allow a certain amount of time to take the exam that is longer than the testing time indicated by the testing software when you begin. This is because Prometric wants you to allow for some time to get settled and take the tutorial about the testing engine.)

The BSCI exam is not an easy exam. You cannot simply read one book and expect to pass it. In fact, the exam is surprisingly difficult; this is so that Cisco can be sure that everyone who passes the test thoroughly understands the subject matter on a conceptual level and is not just good at taking exams. Cisco is very interested in making sure that passing proves that you have the skills to actually implement the features, not just talk about them.

The exam is difficult in subject matter and also in format. You can expect multiple-choice questions, for example, some with multiple answers. You can also expect questions requiring you to pick the correct answer from output screens and configurations. Since 2001, Cisco Systems has also incorporated simulations into their associate-level and professional-level exams. Be prepared to configure a router or perform other internetworking functions while taking the exam.

Another difficult aspect of the exam format is that you are not allowed to go back and change an answer. Candidates who are unsure about the answer will be forced to guess rather than have an extra 15 minutes to think about it at the end of the exam. Those who really know most of the answers will be rewarded by Cisco's attempts to preserve the integrity of the CCNP/CCDP/CCIP

certification. The professional-level certification will mean to everyone that you are highly qualified at the subject at hand.

Although this is a difficult exam, networking professionals can expect to pass if they meet the prerequisites and spend the proper amount of time on training, on-the-job experience, and study. As with most certification exams, you might not pass the first time. Taking the exam a second time, however, is likely to be easier because you have a better idea of what to expect.

There are many questions on the BSCI exam that you might already know through your professional background and experiences, if you meet the prerequisites. This book offers you the opportunity to solidify and build on that knowledge as you make your final preparations to take the BSCI exam. The concepts and commands covered on the exam are not secrets locked in some vault—the information is available in many places and forms, including this book. So, although the exam is difficult, passing is certainly attainable with study.

The Goals of This Book

The primary goal for this book is not only to help you pass the exam, but to ensure that you really understand the concepts and implementation details. The second goal of this book is to provide the most comprehensive coverage of BSCI exam-related topics available, without too much coverage of topics not on the exam. The third and ultimate goal is to get you from where you are today to the point that you can confidently pass the BSCI exam. Therefore, all this book's features, which are outlined in this introduction, are geared toward helping you discover the IP routing topics that are on the BSCI exam, pinpoint where you have a knowledge deficiency in these topics, and determine what you need to know to master these topics.

The Intended Audience

Although the only official prerequisite for CCNP and CCIP certification is CCNA status, and the only prerequisite for CCDP certification is CCDA status, Cisco does not expect you to be able to pass the professional-level exams (such as the BSCI exam) without additional training and experience. This is why Cisco's recommended training for CCNP/CCDP/CCIP involves an official Cisco course. For the routing knowledge required of a CCNP/CCDP/CCIP, Cisco recommends a course called Building Scalable Cisco Internetworks (BSCI) version 2.0.

The BSCI course is targeted toward enterprise network engineers (including systems engineers [SEs], customers, and resellers) who are responsible for network administration and implementation. The targeted audience performs one or more of the following tasks:

- Install and configure network devices
- Design and implement large enterprise networks

- Add services/applications to an existing network, and determine what router configurations are required to support the new services/applications
- Improve traffic flow, reliability, redundancy, and performance throughout the network

NOTE BSCI replaces the old Building Scalable Networks (BSCN) course.

This book is a final-stage preparation tool. Therefore, this book will be most effective as a study resource after you have taken the BSCI course or have acquired an equivalent level of on-the-job experience and training.

The following are the prerequisites for the BSCI course and, for all practical purposes, should be considered prerequisites for using this book effectively:

- Working knowledge of the OSI reference model and the hierarchical model
- Understanding of internetworking fundamentals
- Ability to operate and configure a Cisco IOS device
- Working knowledge of the TCP/IP stack and how to configure a routed protocol such as IP
- Understanding of distance vector routing protocols, such as Routing Information Protocol (RIP) and Interior Gateway Routing Protocol (IGRP), and their operation and configuration
- Ability to determine when to use static and default routes, and how to enable them on a Cisco router
- Ability to display and interpret a routing table from a Cisco router
- Ability to enable a WAN serial connection
- Ability to configure Frame Relay permanent virtual circuits (PVCs) on interfaces and subinterfaces
- Ability to configure an IP standard and extended access list
- Ability to verify router configurations with available tools such as **show** and **debug** commands

NOTE You can acquire these skills from self-paced or instructor-led training sessions and from work experience.

The ideal audience for this book is someone who has attended the Interconnecting Cisco Networking Devices (ICND) course, has achieved CCNA status, and has attended the BSCI course, or who has an equivalent level of on-the-job training and experience with Cisco switches and routers.

Cisco highly recommends that you take courses to support each certification level, but it also recognizes that attending courses might not be an option for everyone. Therefore, if you find yourself struggling with CCNA-level knowledge as you work through this book, consider reviewing a copy of *CCNA Self-Study: Interconnecting Cisco Networking Devices (ICND)*, Second Edition from Cisco Press (December 2003; ISBN 1-58705-142-7). Similarly, if you want routing details at the CCNP/CCDP/CCIP course level, review *CCNP Self-Study: Building Scalable Cisco Internetworks (BSCI)* (ISBN 1-58705-084-6), also from Cisco Press. (Look for the second edition of the BSCI self-study guide in Spring 2004.)

Overview of Cisco Certifications

Cisco's main motivation behind the current certification program is to provide a means of measuring the skills of people working for Cisco resellers and certified partners. Cisco fulfills only a small portion of its orders via direct sale from Cisco; normally, a Cisco reseller is involved. Also, Cisco has not attempted to become the primary source for consulting and implementation services for network deployment using Cisco products, preferring instead to use partners as much as possible. With that business model, there is a great need to distinguish, ensure, and certify the skill levels of the partner companies' employees.

The CCIE program was Cisco's first foray into certifications. Introduced more than ten years ago, the CCIE was designed to be one of the most respected, difficult-to-achieve certifications. To certify, candidates must pass a written test (also given at Prometric), and then Cisco administers a one-day hands-on lab test. Cisco does not publish numbers on pass/fail rates for CCIE or the other certifications, but rumors have the failure rate for all lab test takers at more than 50 percent.

Cisco uses the number of CCIEs on staff as part of the criteria in determining the level of partner status for the company, which in turn dictates the discount received by the reseller when buying from Cisco. (For more insight into reseller certification, go to the Cisco web site at <http://www.cisco.com/en/us/partners/>.) This practice continues to be a good way for Cisco to judge the commitment to having people with proven Cisco skills on staff, which in turn improves customer satisfaction—and customer satisfaction is tied to every Cisco executive's goals.

The CCIE certification became inadequate for helping certify resellers and other partners because, among other factors, the number of partners increased disproportionately to the difficulty of the CCIE exam. Furthermore, many resellers that do not perform services do not require the extreme expertise of a CCIE on staff, other than to receive a better discount. What Cisco needed were certifications that were less rigorous than CCIE and that would allow Cisco more granularity in judging the skills on staff at a partner company. So, Cisco started an entire Cisco Career Certification program, of which CCNP, CCDP, and CCIP are a part.

Cisco developed Routing and Switching career tracks, a Security career track called Cisco Certified Security Professional (CCSP), and several specialization career tracks. Thus far, the Routing and Switching career tracks, which begin with CCNA/CCDA certification, have proven to be the most popular and make up the heart of Cisco certification. The BSCI exam required for CCNP/CCDP/CCIP certification is a central part of the Routing and Switching career tracks.

Four categories of professional-level certifications exist: one to certify implementation skills, a second to certify design skills, a third to certify communications and services skills, and a fourth, introduced in early 2003, to certify security skills. The BSCI is a required exam for three of these four professional-level certifications. Resellers working in a presales environment need more design skills, whereas services companies need more implementation or post-sales skills. The CCNA and CCNP are implementation-oriented certifications, whereas CCDA and CCDP are design-oriented certifications. The CCIP requires a knowledge of the communications and services technologies.

Cisco created two levels of certification in addition to the CCIE: associate level and professional level. The associate level (CCNA/CCDA) is the most basic, and the professional level (CCNP/CCDP/CCIP/CCSP) is the intermediate level between CCNA and CCIE.

Several of the certifications require other certifications as a prerequisite. For instance, CCNP, CCIP, and CCSP certification requires that you have CCNA certification. CCDP requires a CCDA certification. CCIE, however, does not require any other certification before the written and lab tests. CCIE certification is extremely difficult, however, and it is unlikely that someone could achieve that certification without a level of experience and training equaled in attaining and practicing associate- and professional-level certification.

Cisco certifications have taken on a much larger role and importance in the networking industry in recent years. From a career standpoint, Cisco certification can certainly be used to help you get a new job or a promotion. Alternatively, you can have certification added to your performance evaluation plan and then justify a raise based on passing an exam. In addition, not only might passing exams help you land a new job, but it may actually help you make more money.

Exams Required for Certification

To certify for CCNP, CCDP, or CCIP, you must pass multiple exams. This book deals with the BSCI exam 642-801. The qualifying exam for CCIP and CCNP is CCNA, which you can achieve by passing either the CCNA exam #640-801 *or* the INTRO and ICND exams (640-821 and 640-811, respectively). For CCDP, you must first pass the CCDA exam 640-861. The CCNP exams generally match the same topics that are covered in one of the official Cisco courses, but in most cases—and certainly on the BSCI exam—more topics are covered on the exam than are in the course. Table I-1 outlines the exams and the courses with which they are most closely matched.

Table I-1 *Exams and Courses for Which BSCI Is Required, by Certification Level*

Certification	Exam Number	Exam Name	Course Most Closely Matching Exam Requirements
CCNA	640-821	INTRO	Introduction to Cisco Networking Technologies
	640-811	ICND	Interconnecting Cisco Networking Devices
	640-801	CCNA	Interconnecting Cisco Network Devices (ICND)
CCDA	640-861	DESGN	Designing for Cisco Internetwork Solutions
CCNP	642-801	BSCI	Building Scalable Cisco Internetworks (BSCI)
	642-811	BCMSN	Building Cisco Multilayer Switched Networks (BCMSN)
	642-891*	Composite (COMP)	BSCI, BCMSN
	642-821	BCRAN	Building Cisco Remote Access Networks (BCRAN)
	642-831	CIT	Cisco Internetwork Troubleshooting (CIT)
CCDP	642-801	BSCI	Building Scalable Cisco Internetworks
	642-811	BCMSN	BCMSN
	642-871	ARCH	Designing Cisco Network Architectures
CCIP	642-801	BSCI	Building Scalable Cisco Internetworks
	642-641	QoS exam	Implementing Cisco Quality of Service
	640-910	MPLS exam	Implementing Cisco MPLS
	642-661	BGP exam	Configuring BGP on Cisco Routers

* Exam 642-891 meets the same requirements as passing these exams: 642-801 and 642-811. Therefore, you can substitute exam 642-891 for those two exams, but you can expect a longer exam. Exam 642-891 is also the exam to take when recertifying for CCNP and CCDP.

Be cautioned that although the exam coverage and course coverage are similar, there are no guarantees that if you know absolutely everything in the course, you will pass the test. Cisco is moving more toward certifications being tied to technology, not to specific courses. Books in the Cisco Press Exam Certification Guide series can help you prepare for the certification exam, with the added guidance of stressing the most important exam items and coverage of other topics not

taught in the prerequisite courses. Cisco also maintains the right to change the exam content at will to ensure that the exam is current and fair.

Topics on the BSCI Exam

The exact topics that appear on the BSCI exam is a very closely guarded secret. Only those who write the questions for Cisco and who have access to the entire question database truly know what is on the exam. Cisco reveals only general details about the contents and objectives of the BSCI exam. Because Cisco maintains the right to change this information without notice, it is important that you check the web site for the most current information. You can find a list of Cisco exams and the general outline that accompanies each exam at <http://www.cisco.com/go/training>.

Table I-2 lists the BSCI 640-801 exam topics posted on the Cisco web site at the publication time of this book. The table reflects the part of the book in which each topic is discussed. Note that some topics are discussed in more than one part.

Table I-2 Cisco Exam Topics

Part	Cisco Exam Topic
Part I, “IP Routing Fundamentals”	List the key information routers needed to route data. Describe classful and classless routing protocols. Describe link-state router protocol operation. Compare distance vector and link-state routing protocols. Describe concepts relating to extending IP addresses and the use of VLSMs to extend IP addresses. Describe the three-layer hierarchical design model and explain the function of each layer: Access, Distribution, and Core. Given specific requirements, choose the correct routing protocol to meet the requirements. Identify the correct IP addressing scheme, including features of IPv6. Identify the steps to configure a router for Network Address Translation (NAT) with overload, static translations, and route maps. Describe the concepts relating to route summarization and apply them to hypothetical scenarios.

Table I-2 *Cisco Exam Topics (Continued)*

Part	Cisco Exam Topic
Part II, “OSPF”	Describe the features and operation of single-area OSPF. Describe the features and operation of multiple-area OSPF. Given an addressing scheme and other laboratory parameters, identify the steps to configure a single-area OSPF environment and verify proper operation (within described guidelines) of your routers. Given an addressing scheme and other laboratory parameters, identify the steps to configure a multiple-area OSPF environment and verify proper operation (within described guidelines) of your routers. Identify the steps to verify OSPF operation in a single area. Identify the steps to verify OSPF operation in multiple areas. Interpret the output of various show and debug commands to determine the cause of route selection errors and configuration problems.
Part III, “IS-IS”	Explain basic OSI terminology and network layer protocols used in OSI. Identify similarities and differences between Integrated IS-IS and OSPF. List the types of IS-IS routers and their role in IS-IS area design. Describe the hierarchical structure of IS-IS areas. Describe the concept of establishing adjacencies. Given an addressing scheme and other laboratory parameters, identify the steps to configure Cisco routers for proper Integrated IS-IS operation. Identify verification methods that ensure proper operation of Integrated IS-IS on Cisco routers. Interpret the output of various show and debug commands to determine the cause of route selection errors and configuration problems.

(continues)

Table I-2 *Cisco Exam Topics (Continued)*

Part	Cisco Exam Topic
Part IV, “EIGRP”	Describe the features and operation of EIGRP. Given a set of network requirements, identify the steps to configure an EIGRP environment and verify proper operation (within described guidelines) of your routers. Identify the steps to verify EIGRP operation. Interpret the output of various show and debug commands to determine the cause of route selection errors and configuration problems.
Part V, “BGP”	Describe the features and operation of BGP. Explain how BGP policy-based routing functions within an autonomous system. Describe the scalability problems associated with internal BGP. Given a set of network requirements, identify the steps to configure a BGP environment and verify proper operation (within described guidelines) of your routers. Interpret the output of various show and debug commands to determine the cause of route selection errors and configuration problems.
Part VI, “Redistribution and Policy-Based Routing”	Identify the steps to select and configure the different ways to control routing update traffic. Identify the steps to configure policy-based routing using route maps. Identify the steps to configure router redistribution in a network. Explain the use of redistribution between BGP and Interior Gateway Protocols (IGPs). Identify the steps to verify route redistribution. Interpret the output of various show and debug commands to determine the cause of route selection errors and configuration problems.

(continues)

Table I-2 *Cisco Exam Topics (Continued)*

Part	Cisco Exam Topic
Part VII, “Scenarios”	Given specific requirements, choose the correct routing protocol to meet the requirements. Describe the concepts relating to route summarization and apply them to hypothetical scenarios. Given a set of network requirements, identify the steps to configure an EIGRP environment and verify proper operation (within described guidelines) of your routers. Given an addressing scheme and other laboratory parameters, identify the steps to configure a multiple-area OSPF environment and verify proper operation (within described guidelines) of your routers. Identify the steps to configure route redistribution in a network. Identify the steps to select and configure the different ways to control routing update traffic. Describe concepts relating to extending IP addresses and the use of VLSMs to extend IP addresses. Describe the features and operation of EIGRP. Describe the features and operation of multiarea OSPF. Compare classful and classless routing protocols.

Topics in This Book

This section outlines the topics that will be the focus of this book. To accomplish the exam objectives set forth by Cisco (refer to Table I-2), it is important that you have a firm grasp of these topics. The topics correspond to the chapters in which they are covered.

Part I, “IP Routing Fundamentals,” includes the following chapters and topics:

- Chapter 1, “IP Routing Principles”
 - The requirements of the routing process
 - The routing table
 - The differences between a classful and classless routing protocol
 - Routing versus switching

- Chapter 2, “IP Addressing”
 - IP subnetting
 - Variable-length subnet masks (VLSM)
 - Summarization
- Chapter 3, “Designing IP Networks”
 - Criteria in designing IP networks
 - Access lists
 - Methods to reduce network traffic
 - Private addresses
 - Network Address Translation (NAT)
 - IPv6
- Chapter 4, “IP Distance Vector Routing Principles”
 - RIPv1 and RIPv2
 - IGRP and EIGRP
 - Administrative distance
 - Distance vector protocol convergence
 - Interior and exterior gateway protocols
- Chapter 5, “IP Link-State Routing Principles”
 - OSPF
 - IS-IS
 - BGP
 - Link-state routing protocol convergence
 - Distance vector versus link-state routing protocols

Part II, “OSPF,” includes the following chapters and topics:

- Chapter 6, “Using OSPF in a Single Area”
 - OSPF features
 - Electing the designated and backup designated routers
 - OSPF operation in a single area
 - OSPF network topologies

- How OSPF operates in a single NBMA area WAN
- Chapter 7, “Configuring OSPF in a Single Area”
 - How to configure OSPF in a single area
 - How to verify the operation of and troubleshoot an OSPF network
- Chapter 8, “Using OSPF Across Multiple Areas”
 - The issues with interconnecting multiple OSPF areas
 - The differences between the possible types of areas, routers, and LSAs
 - How OSPF operates across multiple areas using NBMA
 - How OSPF supports the use of VLSM and summarization
- Chapter 9, “Configuring OSPF Across Multiple Areas”
 - Required and optional configuration commands for configuring OSPF across multiple areas
 - A working configuration of OSPF across multiple areas
 - Verifying and troubleshooting a multiarea OSPF network

Part III, “IS-IS,” includes the following chapters and topics:

- Chapter 10, “Fundamentals of the Integrated IS-IS Protocol”
 - Introduction to Integrated IS-IS
 - ISO addressing for Integrated IS-IS
 - The Integrated IS-IS hierarchical structure
 - Basic principles of area routing
 - Integrated IS-IS networks and interfaces
 - Network layer protocols used in Integrated IS-IS
- Chapter 11, “Integrated IS-IS Protocol Operation”
 - Integrated IS-IS operation
 - Integrated IS-IS design considerations
- Chapter 12, “Configuring Integrated IS-IS”
 - Configuring Integrated IS-IS
 - Verifying the Integrated IS-IS operation
 - Troubleshooting the Integrated IS-IS operation

Part IV, “EIGRP,” includes the following chapters and topics:

- Chapter 13, “Using EIGRP in Enterprise Networks”
 - The features and operation of EIGRP
 - How EIGRP discovers, chooses, and maintains routes
 - How EIGRP supports summarization
 - How EIGRP functions in an NBMA environment
 - How EIGRP supports large networks
- Chapter 14, “Configuring EIGRP in Enterprise Networks”
 - Configure EIGRP, both in an enterprise network and in an NBMA network
 - Verifying and troubleshooting an EIGRP configuration

Part V, “BGP,” includes the following chapters and topics:

- Chapter 15, “Connecting to Other Autonomous Systems—The Basics of BGP”
 - The features and operation of BGP
 - Design issues with BGP
 - BGP communities, peer groups, and the peering function
 - The configuration of internal and external BGP
 - How to verify the BGP configuration
- Chapter 16, “Implementing and Tuning BGP for Use in Large Networks”
 - Scaling internal BGP
 - Configuring route reflectors
 - Determining policy control using prefix lists
 - Connecting to multiple ISPs
 - Redistributing between interior routing protocols and BGP
 - Configuring and verifying the BGP configuration

Part VI, “Redistribution and Policy-Based Routing,” includes the following chapters and topics:

- Chapter 17, “Implementing Redistribution and Controlling Routing Updates”
 - Selecting and configuring different ways to control routing updates
 - Configuring route redistribution in networks with and without redundant paths between dissimilar routing processes
 - Resolving problems occurring in a redistributed network

- Chapter 18, “Controlling Network Traffic with Route Maps and Policy-Based Routing”
 - Configuring policy-based route maps
 - Verifying and troubleshooting route maps and policy-based routing

Part VII, “Scenarios,” includes the following chapter:

- Chapter 19, “Scenarios for Exam Preparation,” contains three scenarios that test you on various topics covered throughout the book instead of concentrating on a particular technology. This challenges your understanding at a higher level and places the topics in context.

Part VIII, “Appendixes,” includes the following:

- Appendix A, “Answers to Chapter ‘Do I Know This Already?’ Quizzes and Q&A Sections,” provides the answers to the quizzes found in Chapters 1 through 18.
- Appendix B, “Suggested Reading,” provides you with books, standards documents, and web sites that you might find helpful during exam preparation.

The Glossary provides key terms and definitions used throughout this book.

How to Use This Book to Pass the Exam

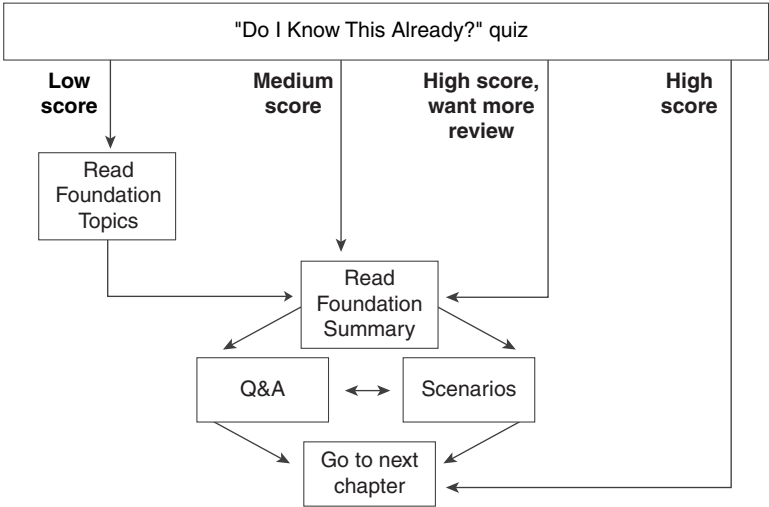
One way to use this book is to start at the beginning and read it cover to cover. Although that would certainly help you prepare, most people do not have that much time to spare, particularly if they already have mastered some of the topics in the book. However, if you want to read the entire book and the review sections on the CD-ROM, and answer all the CD-ROM questions, that is a great way to prepare!

Some of you might want to consider different strategies for how best to use this book, depending on what training and experience you already have. With its prechapter analysis quizzes and chapter-ending summary sections and questions, as well as its traditional foundation sections, this book is designed to help you get the most out of the time you take to study.

The core material for the BSCI exam is covered in Chapters 1 through 18. Figure I-1 shows how to approach the chapter based on your experience and knowledge level.

How you choose to use the study aids in this book might depend on your academic and professional background.

Figure I-1 *How to Use Chapters 1 Through 18*



If you skip to the Foundation Summary, Q&A, and Scenarios sections and have trouble with the material there, you should go back to the Foundation Topics section.

Each core chapter (1 through 18) starts with a “Do I Know This Already?” quiz that allows you to decide how much time you need to devote to studying the subject at hand. Next, the “Foundation Topics” (the core material of the chapter) are presented. This section is the bulk of each chapter. At the end of each chapter, you will find a “Foundation Summary” section that is a collection of tables and quick-reference material that you can use as last-minute review notes. Reviewing each “Foundation Summary” and the “Glossary” toward the end of the book makes for excellent late-stage exam preparation. Each core chapter also has a “Q&A” section of review questions that test you on the chapter’s contents. Finally, where appropriate, some core chapters contain a “Scenarios” section that tests you further on the material.

The appendixes contain materials for your reference. Appendix A contains the answers to each chapter’s “Do I Know This Already?” and “Q&A” quizzes. The answers to the “Scenarios” questions can be found at the end of each chapter.

This book is also accompanied by a CD-ROM that offers multiple-choice questions based on the entire book’s content. Each question in the CD-ROM refers you to the chapter and section it is drawn from.

Command Syntax Conventions

The conventions used to present command syntax in this book are the same conventions used in the *Cisco IOS Command Reference*, as follows:

- **Boldface** indicates commands and keywords that are entered literally as shown. In examples (not syntax), boldface indicates user input (for example, a **show** command).
- *Italics* indicates arguments for which you supply values.
- Square brackets ([and]) indicate optional elements.
- Braces ({ and }) contain a choice of required keywords.
- Vertical bars (|) separate alternative, mutually exclusive elements.
- Braces and vertical bars within square brackets—for example, [x {y | z}]—indicate a required choice within an optional element. You do not need to enter what is in the brackets, but if you do, you have some required choices in the braces.

CAUTION Throughout this book, examples, sample configurations, and figures use both private and valid IP addresses. These addresses are used for educational purposes only and are not intended for use by readers while connected to the Internet. Because it is possible that some of these addresses have been assigned to organizations, any connection, however innocent, will be construed as unauthorized access or Internet trespass.

It is important in all cases that any configurations are performed in a controlled lab environment that is not connected to either your organization's network or the Internet. This is not merely because the addresses used in the book may be real addresses, but because it would be foolhardy to practice your networking skills in an operational environment.

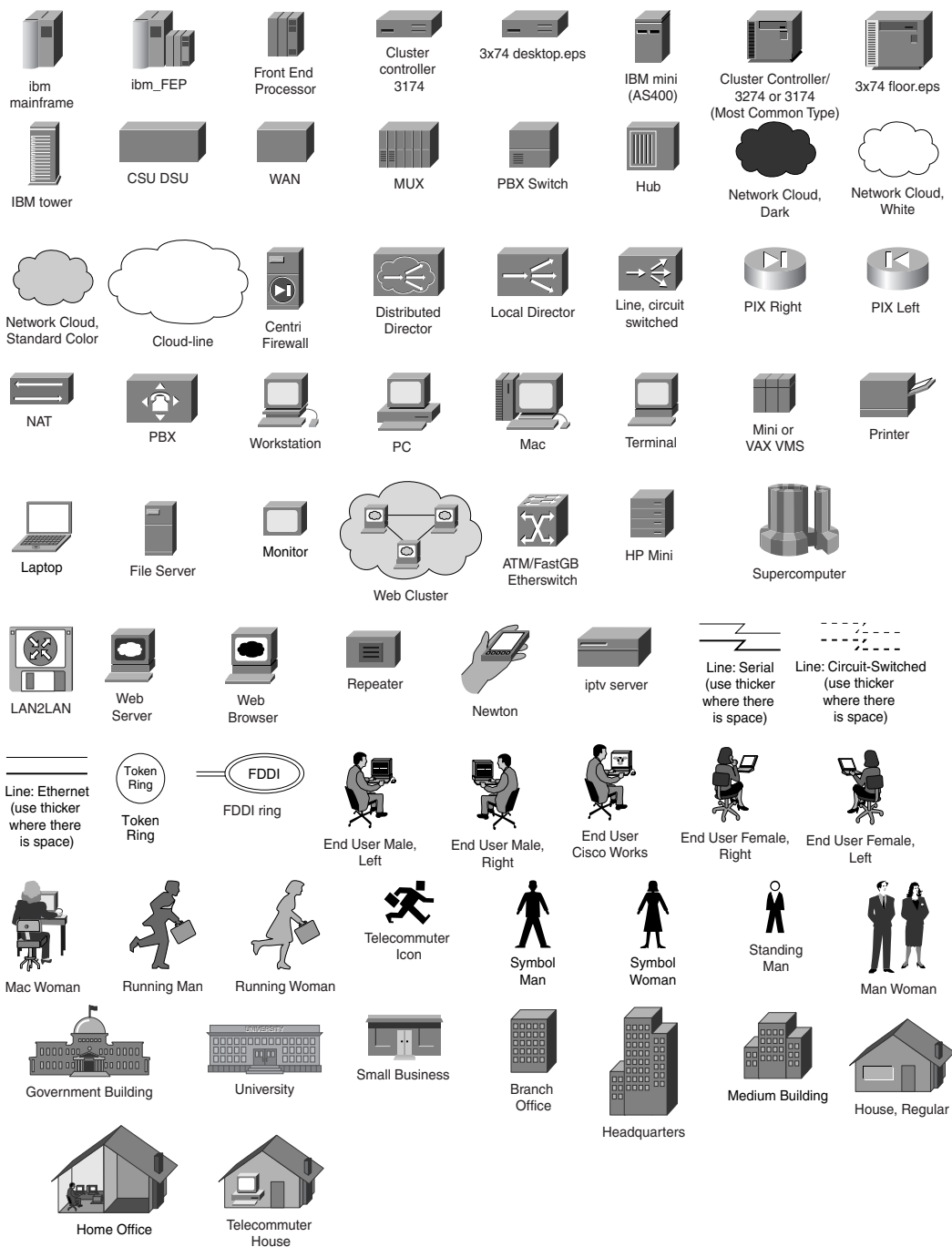
Neither Cisco, Cisco Press, nor its staff is responsible for any action taken by the readers of this book in conjunction with the unauthorized use of IP addresses.

Icons Used in This Book

Throughout this book, you will see the following icons used for networking devices:



The following icons are used for peripherals and other devices:



The following icons are used for networks and network connections:



Strategies for Exam Preparation

There are basically five different categories of BSCI exam candidate:

- Those who have taken the BSCN/BSCI course
- Those who have attended the Cisco Networking Academies
- Those who will not be taking any classes and have not had much experience
- Those who will not be taking any classes but have some experience

I've Taken BSCI—Now What?

First, let me say that you've taken the best path to prepare yourself. Let me temper that with the fact that if you retain more than 50 percent of what you heard in class, you are an extraordinary person. That said, in my opinion, you need to follow these strategies:

- **Strategy 1** — Use this book exactly as described in the opening pages of Chapters 1 through 18. Each of the core chapters begins with a quiz that helps you assess what you need to study. It then directs you to the appropriate sections in the chapter, rather than requiring you to read the entirety of each chapter.
- **Strategy 2** — Use the directions at the beginning of Chapter 19 to direct your final study before the exam. Chapter 19 is designed to review many concepts; in addition, it outlines a good process for study in the days leading up to your exam.

By using these strategies, you will fill in your gaps in knowledge and will be confident taking your BSCI exam.

I've Taken the Cisco Networking Academy Courses—Now What?

First, I'll start by congratulating you on having the foresight to get into the Cisco Networking Academy program. For those of you who did not take the Cisco Networking Academy track and are wondering what it is, visit <http://www.cisco.com/en/US/learning/index.html> and click to the Cisco Networking Academy page. The Networking Academy curriculum does a great job of preparing you with the skills and knowledge that you need to pass the BSCI exam. Unfortunately, your study was probably spread over several semesters, and possibly over a couple years. So, you might have forgotten the details that you do not use frequently. On to the strategies for success on CCNP/CCDP/CCIP—and, in particular, the BSCI exam:

- **Strategy 1** — Pull out your Networking Academy curriculum and notes and reread them. Most people find that their memory is exercised better by seeing familiar material, and even more so when they wrote it down themselves.

- **Strategy 2** — Use this book exactly as described in the opening pages of Chapters 1 through 18. Each of the core chapters begins with a quiz that helps you assess what you need to study. It then directs you to the appropriate sections in the chapter, rather than requiring you to read the entirety of each chapter.
- **Strategy 3** — Make it a point to read the sections in this book that cover some of the conceptual topics and standards. This focus will be helpful because the Networking Academy is more oriented toward building skills, not imparting theoretical knowledge.
- **Strategy 4** — Use the directions at the beginning of Chapter 19 to direct your final study before the exam. Chapter 19 is designed to review many concepts; in addition, it outlines a good process for study in the days leading up to your exam.

I'm New to Internetworking with Cisco, and I Will Not Be Taking the BSCI Course—Now What?

You can take and pass the BSCI exam without taking any courses. Cisco wants you to take the recommended courses for all the exams, though. The Cisco motivation is that the more people who understand Cisco products, ultimately the happier the customers will be and the more efficiently the networks will run, which is in everyone's interests. In addition, Cisco believes that its official training is the best way to teach people about its products.

If you are not taking the course, you will need more than just this book to prepare. Here are my strategy suggestions for your case:

- **Strategy 1** — Read *CCNP Self-Study: Building Scalable Cisco Internetworks (BSCI)* by Cisco Press. Although BSCI is not entirely a course-based test, the BSCI course is listed as the recommended course for that exam. In *CCNP Self-Study: Building Scalable Cisco Internetworks (BSCI)*, the basic text and illustrations come from the course material but are expanded and reorganized to work well in a self-study format.
- **Strategy 2** — After reading *CCNP Self-Study: Building Scalable Cisco Internetworks (BSCI)*, use this book exactly as described in the opening pages of Chapters 1 through 18. Each of the core chapters begins with a quiz that helps you assess what you need to study. It then directs you to the appropriate sections in the chapter, rather than requiring you to read the entirety of each chapter.
- **Strategy 3** — Use the guidelines at the beginning of Chapter 19 to direct your final study before the exam. Chapter 19 is designed to review many concepts; in addition, it outlines a good process for study in the days leading up to your exam.

I've Learned a Lot About CCNP Topics Through Experience, but I Will Not Be Taking the BSCI Course—Now What?

If you feel like you know a fair amount about professional-level routing topics already (at a level that makes taking the BSCI course not very worthwhile), but you are worried about the few topics that you simply have not worked with, then the following strategies are for you:

- **Strategy 1** — Use this book exactly as described in the opening pages of Chapters 1 through 18. Each of the core chapters begins with a quiz that helps you assess what you need to study. It then directs you to the appropriate sections in the chapter, rather than requiring you to read the entirety of each chapter.
- **Strategy 2** — Use the guidelines at the beginning of Chapter 19 to direct your final study before the exam. Chapter 19 is designed to review many concepts; in addition, it outlines a good process for study in the days leading up to your exam.

Strategies for the Exam Day

Here is a reminder of some simple things you can do to help prepare yourself for the day of the exam.

On the day before the exam:

- Call Pearson VUE or Prometric to confirm your seat, exam time, and location of the exam center. Also check the confirmation number that was allocated for your exam.
- Ensure that you have directions for the center and the location of the nearest parking garage.
- Have a relaxing evening; do not be tempted to heavily review because this will simply emotionally exhaust you and prevent a good night's sleep. If you cannot resist some studying, simply read through the Q&A sections in Appendix A.

On the exam day:

- Eat a nutritious meal before you leave. Rumbling stomachs are distracting, and it is proven that your brain functions better when fueled.
- Leave plenty of time to get to the testing center, park, and have a few moments to relax before the exam. Allow at least half an hour for traffic jams and the like.
- The testing center will provide pen and paper. You are not allowed anything in the exam room, except a refreshment and the pen and paper provided. Leave all those heavy books at home.
- Wear loose, comfortable clothing.

During the exam:

- Work out the timing. If the exam still has 61 questions and you allowed 75 minutes to complete, that means you have approximately 1.25 minutes per question. Because you cannot return to questions after you have passed them, you must try to allocate an equal amount of time to each

question. It is counterproductive to miss questions or to guess unless you really have no idea of the answer. The exam tells you how many questions are left to answer and how much time is left.

- If you do not know the answer to a question, try answering the question by a process of elimination. As Sherlock Holmes said, “When you have eliminated the impossible, whatever remains, however improbable, must be the truth.”
- Use the paper provided to work out the logic of some questions.
- Try to stay calm. Remember that the exam can be taken multiple times, so even if you are struggling, you can use the exam to your advantage by remembering what topics are causing you trouble.

Conclusion

The Cisco certification has great value in the networking environment. It proves your competence and dedication. It is required for several certifications, and it is a huge step in distinguishing yourself as someone who has proven knowledge of Cisco products and technology, instead of just claiming that you know it.

The *CCNP BSCI Exam Certification Guide* is designed to help you attain CCNP, CCDP, or CCIP certification. It is a certification book from the only Cisco-authorized publisher. We at Cisco Press believe that this book will help you achieve certification, but the real work is up to you. We hope you find your time well-spent with this book. Good luck!

Part I: IP Routing Fundamentals

Chapter 1 IP Routing Principles

Chapter 2 IP Addressing

Chapter 3 Designing IP Networks

Chapter 4 IP Distance Vector Routing Principles

Chapter 5 IP Link-State Routing Principles

Part I covers the following Cisco BSCI exam topics:

- List the key information that routers need to route data
- Describe classful and classless routing protocols
- Describe link-state router protocol operation
- Compare distance vector and link-state routing protocols
- Describe concepts relating to extending IP addresses and the use of VLSMs to extend IP addresses
- Describe the three-layer hierarchical design model and explain the function of each layer: Access, Distribution, and Core
- Given specific requirements, choose the correct routing protocol to meet the requirements
- Identify the correct IP addressing scheme, including features of IPv6
- Identify the steps to configure a router for Network Address Translation with overload, static translations, and route maps
- Describe the concepts relating to route summarization and apply them to hypothetical scenarios



This chapter covers the following topics, which you need to understand to pass the CCNP/CCDP/CCIP BSCI exam:

- Understanding routing protocol fundamentals
- Types of routing protocols
- The routing table
- Routing versus switching

IP Routing Principles

The topics in this chapter are the basis of the BSCI course that feeds the BSCI exam. It is therefore the first chapter of the book, as all remaining chapters build from the subjects discussed here.

This chapter describes the different types of routing protocols for IP, explains how to read the routing table and keep it current, and compares the routing and switching functions.

The concepts of routing with IP and the mechanics of the routing process are dealt with generically in this chapter as a foundation for the subsequent chapters, which deal with the individual routing protocols.

The topics presented here will directly reflect questions on the BSCI exam. If you do not understand the contents of this chapter, it will be impossible for you to pass the exam. The subsequent chapters assume the comprehension of the subjects covered in this chapter.

“Do I Know This Already?” Quiz

The purpose of the “Do I Know This Already?” quiz is to help you decide what parts of this chapter to use. If you already intend to read the entire chapter, you do not necessarily need to answer these questions now.

The 12-question quiz, derived from the major sections in the “Foundation Topics” portion of the chapter, helps you to determine how to spend your limited study time.

Table 1-1 outlines the major topics discussed in this chapter and the “Do I Know This Already?” quiz questions that correspond to those topics.

Table 1-1 “Do I Know This Already?” Foundation Topics Section-to-Question Mapping

Foundation Topics Section	Questions Covered in This Section
Understanding Routing Protocol Fundamentals	1–3
Types of Routing Protocols	4–6

(continues)

Table 1-1 “Do I Know This Already?” Foundation Topics Section-to-Question Mapping (Continued)

Foundation Topics Section	Questions Covered in This Section
The Routing Table	7–9
Routing Versus Switching	10–12

NOTE The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark this question wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

1. Which of the following criteria would select a route as the preferred path to a destination network to be added to the routing table?
 - a. The route is the only available path to the destination network.
 - b. The route has one interface configured for IP.
 - c. The route has the lowest administrative distance.
 - d. The route has the lowest metric.

2. Which of the following best describes a routing protocol?
 - a. A set of rules that describes how Layer 3 routing devices will send updates between each other about available networks and build a routing table
 - b. The mechanism by which a router describes the autonomous system
 - c. The means by which a router learns about the Internet
 - d. The rules that state how a router builds a routing table

3. What is the forwarding process?
 - a. The means by which a datagram is selected to be in the routing table
 - b. The process that determines which path to select from the routing table
 - c. How networks are forwarded between autonomous systems
 - d. The switching process using the memory cache

4. What is the purpose of the command **ip classless** ?
 - a. To force a classful routing protocol to send the subnet mask in its routing updates
 - b. To allow VLSM to be configured on directly connected interfaces

- c. To allow the redistribution of OSPF subnets into RIPv1
 - d. To allow the router to ignore the classful boundaries of the networks in its routing table and simply route to the longest prefix match it can find
5. Which of the following statements characterizes classless routing protocol?
- a. Within the same IANA classful network address, subnets are exchanged between the routers without the subnet mask.
 - b. Router interfaces using the same IANA classful network address can have different subnet masks.
 - c. IGRP and EIGRP are the only distance vector routing protocols that can be configured as classless routing protocols.
 - d. Within the same IANA classful network address, subnets are exchanged between the routers.
6. Which of the following two statements is a characteristic of a classful routing protocol?
- a. Routes exchanged between foreign networks are summarized to the IANA classful address.
 - b. Within the same network (IANA network number), different subnet masks can be configured.
 - c. All the interfaces on all the routers within an IANA network number must share the same subnet mask.
 - d. All interfaces on a router must share the same subnet mask.
7. Which of the following is an instance of when a static route is used?
- a. On a slow speed dialup link
 - b. On networks where surge protectors are not in use
 - c. When there are many links to the Internet
 - d. On unstable networks
8. What does 0.0.0.0 signify in an IP routing table?
- a. Indicates the router whose routing table is being displayed
 - b. Indicates that the router has no directly connected interfaces
 - c. Indicates that no routing protocol is running
 - d. Shows the default route for the routing table

9. What is the command syntax to empty the Cisco routing table of all its routes?
- a. **clear ip route ***
 - b. **clear ip route table**
 - c. **reset ip route *.*.*.***
 - d. **clear ip route 0.0.0.0**
10. Cisco makes a distinction between the routing function and the switching function of a router. Which of the following best describes the routing function?
- a. Moving data to the end destination that is in cache
 - b. Load balancing
 - c. Moving the datagram across the router
 - d. Path determination
11. What information is found in the routing table in order to process switch a packet?
- a. The address of the neighboring router that will forward the traffic
 - b. The MAC address of the remote network
 - c. The destination address of a static route
 - d. The path with the lowest administrative distance
12. Which of the following questions are asked by the routing function when trying to route a datagram?
- a. If the destination network is not in the routing table, is there a default network configured?
 - b. Is fast switching enabled?
 - c. Which is the best path to that remote network?
 - d. Are there multiple equal-cost paths?

The answers to the “Do I Know This Already?” quiz are found in Appendix A, “Answers to Chapter ‘Do I Know This Already?’ Quizzes and Q&A Sections.” The suggested choices for your next step are as follows:

- 6 or less overall score—Read the entire chapter. This includes the “Foundation Topics” and “Foundation Summary” sections, the “Q&A” section, and the “Scenarios” at the end of the chapter.
- 7–9 overall score—Begin with the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. If you have trouble with these exercises, read the appropriate sections in “Foundation Topics.”
- 10 or more overall score—If you want more review on these topics, skip to the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. Otherwise, move to the next chapter.

Foundation Topics

Understanding Routing Protocol Fundamentals

This section covers the definition, purpose, and operation of a routing protocol. It also covers the difference between a routing protocol and a routed protocol. You need to understand the definition of a protocol to understand exactly what a routing protocol is trying to achieve.

The Definition of a Routing Protocol

In simple terms, a *protocol* is an agreed upon set of rules that determines how something will operate. A *routing protocol* is a set of rules that describes how Layer 3 routing devices send updates between each other about the available networks. If more than one path to the remote network exists, the protocol also determines how the best path or route is selected.

A routing protocol is the mechanism used to update the Layer 3 routing devices. When they all have the same accurate understanding of the network, they can route the data across the best path.

How the Routing Protocol Works

Participating routers advertise the routes that they know about to their neighbors in routing updates. Routes learned from routing updates are dynamic routes held in the routing table. (You will learn more about routing tables later in this chapter in the section “The Routing Table.”)

The routing process is confusing until you realize that there are actually three steps involved in building, maintaining, and using the routing table. These three steps are independent of one another and include the following:

1. The routing protocol sends the information about the routes or networks within the autonomous system, such as RIPv1, IGRP, and EIGRP, and between autonomous systems with BGP-4.
2. The routing table receives updates from the routing protocol and provides the forwarding process with information on request
3. The forwarding process determines which path to select from the routing table in order to forward a datagram.

These three steps use the following criteria to make decisions:

- Metrics—The routing protocol uses metrics to calculate which path is the best path to the remote destination network. Multiple IP routing protocols cannot easily share information because their metrics are completely different.
- Administrative distance—If more than one routing process is running on the router, the administrative distance is used to select which protocol will update the routing table. This is based on which routing protocol is considered the most reliable source of accurate information.
- Prefix length—The forwarding process will use the route where the most number of subnet bits match that of the destination network. It chooses the most specific match, known as the match to the longest prefix length.

Routing and Routed

Distinguishing between the datagram and the routing protocol used to determine the path of the datagram is important. The distinction is between the *routed* protocol and the *routing* protocols.

The routed protocol is the Layer 3 protocol used to transfer data from one end device to another across the network. The routed protocol is the Layer 3 datagram that carries the application data in addition to the upper-layer information.

The routing protocol is the protocol used to send updates between the routers about the networks that exist in the organization, thereby allowing the routing process to determine the path of the datagram across the network.

Table 1-2 provides a list of routed protocols and their corresponding interior routing protocols.

Table 1-2 Routing and Routed Protocols

Routed Protocol	Corresponding Interior Routing Protocol ^a
AppleTalk	RTMP, AURP, EIGRP
IPX	RIP, NLSP, EIGRP
Vines	RTP
DECnet IV	DECnet
IP	RIPv1, RIPv2, OSPF, IS-IS, IGRP, EIGRP

a. IGRP and EIGRP are Cisco Systems proprietary routing protocols.

Types of Routing Protocols

Although the switching and routing functions within the router are set, there are many differences among the different routing protocols.

The routing protocols are essentially applications on the router. Their purpose is to ensure the correct and timely exchange of information about the network between the routers so that the routers can successfully perform the routing and switching functions described previously.

IP routing protocols can be divided into several distinct groups. The first difference is between protocols that send the subnet mask in the updates and the older protocols that do not. These are labeled classless and classful protocols, respectively.

Classful Routing

Classful routing protocols do not carry the subnet or routing mask in the update. The older distance vector protocols tend to be classful. This incapability of carrying the subnetting information leads to design constraints in the IP network.

Classful IP routing protocols include RIPv1 and IGRP. The characteristics of a classful routing protocol are as follows:

- Summarization occurs at the network boundary.
- Routes exchanged between foreign networks are summarized to the IANA classful network boundary.
- Within the same network (IANA classful network), subnet routes are exchanged by routers, without the mask.
- The subnet mask is assumed to be consistent for a IANA classful network used within a network, so all router interfaces must share the subnet mask for interfaces in the same IANA classful network.

The rules of classful addressing lead the router to make specific decisions. Therefore, if there is an entry in the routing table for a specific subnet, the datagram is forwarded to that destination. If the destination subnet is unknown, the datagram is dropped. There is nothing unexpected here. If there is a default network configured for the router, you might expect it to be used. However, the default network is only used in classful routing if there is no knowledge of the network at any level. Thus, if the major network is known, the datagram is dropped even if there is a default network.

The forwarding decision made by a router using classful forwarding is as follows:

1. Send the datagram to the subnet if there is an entry in the routing table.
2. If there is no entry in the routing table, discard the datagram.

3. If there is an entry for a major network but not for the specific subnet, discard the datagram.
4. If there is an entry for a major network but not the specific subnet, do not consider the default network but instead discard the datagram.
5. If there is no entry for either the major network or the subnet, but there is a default route, forward the datagram to the default network.

Classless Routing

Classless routing protocols were designed to overcome the constraints of classful routing, which include the following:

- The utilization of address space might be inefficient.
- Variable-length subnet mask (VLSM) is not possible within the network.
- The inability to use VLSM can lead to very large routing tables and overloaded network links, saturated with routing updates. The routing protocols that are capable of classless routing are OSPF, EIGRP, RIPv2, IS-IS, and BGP.

The characteristics of a classless routing protocol are as follows:

- Router interfaces within the same network can have different subnet masks (VLSM).
- The classless routing protocols support the use of classless interdomain routing (CIDR).
- Some routes can be summarized within the major IANA classful network number. This is done manually.

To utilize the advantages of classless routing, Cisco created the command **ip classless**, which is the default configuration in current IOS releases. Although this command might seem to solve many of your problems, you need to consider a few caveats. The following section describes the use of the **ip classless** command and, in particular, how it changes the decisions made by the routing process.

The ip classless Command

The **ip classless** command changes the forwarding decisions made on entries in the routing table. It does not change the way the routing table is built, but rather how datagrams are routed based on the contents of the routing table.

The **ip classless** command is the default configuration from IOS software 12.0. To turn off this facility, use the global configuration command **no ip classless**. Issuing this command will create the actions described in the previous list.

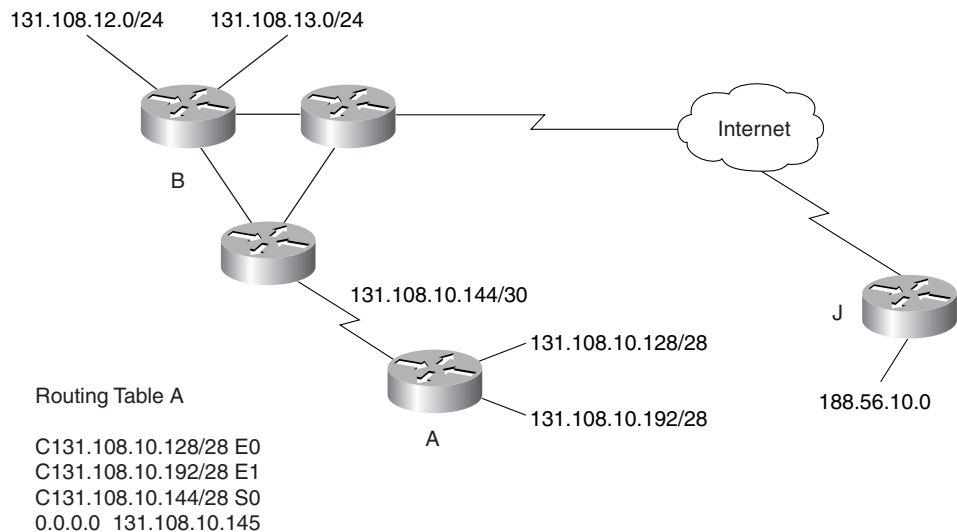
Although it makes sense that Cisco would create the inverse of any command that it introduces, there are instances when making classless routing decisions can cause problems. In these instances, the command **no ip classless** should be issued.

Consider a network using a Class B addressing system of 131.108.0.0. A remote stub router is configured; it has two Ethernet connections and a serial connection to the hub. The stub router is configured with a default route and no routing protocol running. If the router is running an IOS software version earlier than 12.0, it would be able to connect to the Internet, but not to any other network within its classful (or major) network.

NOTE Default routes learned through IS-IS or OSPF ignore the **no ip classless** command.

In Figure 1-1, Router A is a stub router running Cisco IOS version 11.0 and has not had IP classless configured on it. Although it can reach the network—188.56.10.0 on Router J—on the other side of the Internet through the default network, it cannot see or reach the network within its own autonomous system, 131.108.12.0/24, because it has a different subnet mask.

Figure 1-1 The `ip classless` Command



The Routing Table

The previous sections discussed how routing decisions are made and referenced the routing table. This section demystifies the routing table by dealing in depth with the following topics:

- The routing table fields
- How to keep the routing table current and correct
- Other methods of entering routes into the routing table

The Routing Table Fields

The router will reference the routing table and make a decision about forwarding data datagrams to the end destination identified in the destination address of the datagram/packet.

Table 1-3 shows the fields that are present in a typical routing table.

Table 1-3 The Routing Table

Network	Outgoing Interface	Metric	Next Hop
140.100.100.0 /24	E0	6	131.108.13.15
140.100.110.0 /24	E0	7	131.108.13.15
140.100.120.0 /24	E0	8	131.108.13.15
140.100.130.0 /24	E0	8	131.108.13.15
166.99.0.0 /16	E1	10	131.108.14.11
166.90.0.0 /16	E1	11	131.108.14.11
145.0.88.0 /24	S0	3	131.108.10.9

It is useful to look at each field in the routing table to determine the functionality of the table to the routing process. The next sections cover the following fields of the routing table:

- The Network field
- The Outgoing Interface field
- The Metric field
- The Next Hop field

The Network Field

The Network field contains the networks that the router knows exist in the organization. These entries either were entered manually as *static routes* or were learned via a routing protocol as *dynamic routes*. *Default routes* can be learned either by manual configuration or by routing updates.

When a datagram comes into the router, the routing process attempts to forward it to the remote network, where it should find the destination host. To successfully forward the datagram to the destination host, the routing process must know that the remote network exists. The routing process determines the remote network’s existence by looking in the routing table for the remote network.

Typically, only the network portion of the address is stored in the table. Using the hierarchical strength of the addressing keeps the routing table small and the lookup short. The routing process

makes a decision based on the longest match to find the most specific entry. This ensures that if VLSM has been deployed, the most specific network is chosen. Cisco IOS code mandates that the longest match can be a /32 or 255.255.255.255 mask. This is a match based on the full host address and is used in specific situations, such as an OSPF environment. It is not encouraged as a common configuration because the size of the routing table grows rapidly. VLSM is described in detail in Chapter 2, “IP Addressing.”

The routes in the table are held in an order that speeds up the lookup process, ensuring that the routing decision is streamlined.

Later in the chapter, in the section “How to Keep the Routing Table Current and Correct,” you will see how the networks are placed in the table and how path selection to a remote network is chosen.

The Outgoing Interface Field

The Outgoing Interface field in the routing table indicates the following:

- To which interface to send the datagram
- Through which interface the routing update came

The Outgoing Interface field corresponds to the column in the routing table that stores the information about the interface on the router to which the routing process sends the datagram. This is the first step of its journey, the exit point of the router.

The routing process must know which interface queue to use to send the outbound datagram. The outgoing interface field also informs the administrator of the interface through which the network was heard in the routing update, or, more accurately, the interface through which the chosen network was heard.

The Metric Field

The *metric* is a value that is assigned to each path based on the criteria specified in the routing protocol. The Metric field is used to determine the best path to use if there are multiple paths to the remote destination network. The metric used depends on the routing protocol.

This value is used to choose between different paths to the same destination network to select the best path. If the values are the same, the router either selects the path that it heard first or uses all of the multiple paths, sending the datagrams across each route.

Table 1-4 shows how the metrics are derived by the different routing protocols.

Table 1-4 *Routing Protocol Metrics*

Routing Protocol	Metric
RIPv1	Hop count
RIPv2	Hop count
IGRP	Bandwidth, delay, load, reliability
EIGRP	Bandwidth, delay, load, reliability
OSPF	Cost (Although the standards do not specify how to calculate the cost, the Cisco default states that the cost of an interface is inversely proportional to the bandwidth of that interface. A higher bandwidth indicates a lower cost, making higher speed links preferable.)
IS-IS	Cost (Although the standards do not specify how to calculate the cost, the Cisco default states the cost is 10 regardless of bandwidth.)

NOTE By default, on a Cisco router, if multiple equal-cost paths exist in IP, up to six paths are used in a round-robin manner to load balance the traffic across the network.

The Next Hop Field

The *next hop* is the destination address of the next forwarding router. The address of the next hop will be on the same subnet as the outgoing interface. However, there are exceptions to this rule, for example, the next hop in Internal BGP. Internal BGP is beyond the scope of this chapter but is described in more detail in Chapter 15, “Connecting to Other Autonomous Systems—The Basics of BGP-4.”

The purpose of identifying the next hop is so that the router can create the Layer 2 frame with the destination address. When troubleshooting, remember that the next hop address is the address of the router directly connected to the forwarding router. Therefore, the address of the next hop shares the same subnet as the determining router.

Now that you understand the fields of the routing table, the next section discusses how to keep the routing table current.

How to Keep the Routing Table Current and Correct

The routing table in *every* router within the autonomous system must be both accurate and up-to-date so that datagrams can be directed across the network to their destination. The more recent routing protocols are more efficient than the early offerings of RIP, using fewer network resources,

which allows networks to scale while maintaining routing tables that are accurate. For example, the older routing protocol RIP sends out the entire routing table every 30 seconds to a broadcast address. On the other hand, OSPF updates contain information only about network changes and are sent only when that change occurs, using multicast addressing. To ensure that the routing tables are synchronized, OSPF sends a summarization of the entire routing table 30 minutes after the last update was sent.

The accuracy of the table will be affected by how quickly it responds to changes in the network. These changes include the following:

- Learning new networks
- Learning a better path to an existing network
- Learning that a network is no longer available
- Learning an alternative route to a network

How each of these changes is achieved depends on the routing protocol.

Emptying the contents of the routing table and thus forcing the router to relearn the information about the network is very useful in troubleshooting a network.

This command empties all the routes from the table:

```
Router# clear ip route *
```

This command removes the specific network from the table:

```
Router# clear ip route {network [mask] / *} }
```

The following section gives an example of a routing table. In the BSCI exam, you might be asked to interpret the output of the **show ip route** command. Therefore, you will need to be able to extrapolate information from this table. The following section shows you how.

The show ip route Command

The **show ip route** command is used to show the IP routing table on the router. It details the network as known to the router and its sources for the information (such as the routing protocols). This command is excellent for troubleshooting configuration errors and understanding how the network is communicating routing information.

Example 1-1 shows the output of the **show ip route** command. Table 1-5 explains how to read this information.

Example 1-1 *The show ip route Command Output*

```
SanJose#show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, * - candidate default
       U - per-user static route, o - ODR
       T - traffic engineered route

Gateway of last resort is not set

      140.100.0.0/28 is subnetted, 3 subnets
C       140.100.17.192 is directly connected, FastEthernet3/0
C       140.100.17.128 is directly connected, FastEthernet1/0
C       140.100.32.0 is directly connected, Fddi2/0

Bldg_1#show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, * - candidate default
       U - per-user static route, o - ODR

Gateway of last resort is not set

      140.100.0.0/28 is subnetted, 3 subnets
O       140.100.17.192 [110/20] via 140.100.17.129, 00:07:44, Ethernet0
C       140.100.17.128 is directly connected, Ethernet0
O       140.100.32.0 [110/11] via 140.100.17.129, 00:07:44, Ethernet0
```

Table 1-5 *Explanation of the **show ip route** Command That Was Performed on Router Building 1*

Field	Explanation
O	Indicates the protocol that derived the route. Possible values include the following: I —IGRP-derived R —RIP-derived O —OSPF-derived C —Connected S —Static E —EGP-derived B —BGP-derived i —IS-IS-derived D —EIGRP
140.100.17.192	Indicates the remote network.
[110/20]	The first number in the brackets is the administrative distance of the information source; the second number is the metric for the route.
via 140.100.17.129	Specifies the address of the next router to the remote network.
00:07:44	Specifies the last time that the route was updated in hours:minutes:seconds.
Ethernet0	Specifies the interface through which the specified network can be reached.

The **show ip route** command is useful to verify that the configuration has worked and that the network is functioning correctly.

Other Methods of Entering Routes into the Routing Table

Using a routing protocol is the easiest method to create and maintain an accurate and current routing table. However, it is not the only way, nor is it always the most efficient way to inform the router of the different routes available in the autonomous system. If the router has few resources and is connected to a stub network, it might be more efficient to simply state the path to a router that has full knowledge of the network. This section considers some of the other methods—including static routes, default static routes, On Demand Routing, and floating static routes—and how, when, and why you might use them.

Static Routes

The routing process sends and receives updates to and from other routers in the autonomous system. These updates populate the routing table, by which forwarding decisions are made and datagrams are sent to the destination address.

There are instances, however, when it is better for the administrator to manually insert a route into the routing table, thus either bypassing the routing process altogether and having no routing protocol running on the router, or adding individual routes into the routing table.

Manually configuring the routing table means adding static routes into the routing table. The advantages of static routes are the conservation of both router and network resources. A routing protocol requires memory, CPU, and bandwidth not simply of the immediate router but of every router in the autonomous system.

The trade-off is that the resources required from the network administrator are considerable. The administrator is responsible for adding every static route. If there is a change in the network topology, it is the administrator's responsibility to change every affected static route within the network. By definition, these routes cannot dynamically correct themselves. Therefore, the network will not converge until the routers are reconfigured.

It seems unlikely that there could be any situation that could warrant such a labor-intensive solution, but there are a few. These include

- Links have very low bandwidth, such as dialup links.
- The administrator needs control over the link.
- The link is a backup to the dynamically learned route.
- There is only one path to the remote network, such as a stub network.
- The router has very limited resources and cannot run a routing protocol.
- The administrator needs to control the routing table to allow a classful and classless routing protocol to populate the routing table. Classful and classless routing was explained earlier in this chapter.

The command for configuring a static route is a global configuration command. The correct syntax is as follows:

```
ip route prefix mask {ip-address | interface-type interface-number} [distance] [tag tag]  
[permanent]
```


Table 1-6 describes the usage and meaning of the command parameters.

Table 1-6 *The ip route Command Description*

Syntax	Description
<i>prefix</i>	IP route prefix for the destination
<i>mask</i>	Prefix mask for the destination
<i>ip-address</i>	IP address of the next hop that can be used to reach that network
<i>interface-type interface-number</i>	Network interface type and interface number
<i>distance</i>	(Optional) An administrative distance
tag tag	(Optional) Tag value that can be used as a “match” value for controlling redistribution via route maps
permanent	(Optional) Specifies that the route will not be removed, even if the interface shuts down

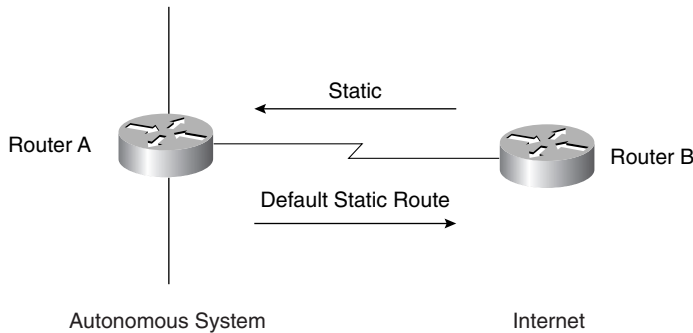
Default Static Routes

On occasion, you need to use a static route to create a default route. A *default route* is the route that is used if there is no entry for the specific destination route in the routing table. If a datagram comes in destined for the Internet, the routing table, unaware of all the networks outside its autonomous system, will direct the traffic to the edge router that connects to the outside world.

Occasions for using a default route include:

- Connecting to the autonomous system from a stub network
- Connecting to the Internet

Figure 1-2 shows an example of when to use a default route and static route together. Router A is the edge router in the autonomous system and has a default route to the Internet. This default route is propagated to every router in the organization. If a router within the organization is asked to route to a network that is not in its routing table, it directs the traffic to Router A and to the Internet. Likewise, Router B does not need the entire routing table of the organization to which Router A belongs; a static route is configured with the IANA classful network for the organization. This ensures that the routing tables are no larger than they need to be and that the link does not have to suffer the overhead of a routing protocol.

Figure 1-2 *Static and Default Static Routes*

NOTE Remember that when a static route is configured, you have simply told the connected router how to get to the remote network. If routing is disabled on that remote router, it cannot reply. A reciprocal route (static or dynamic) is required.

Also, static routes that point to an interface are fine for point-to-point circuits, but in point-to-multipoint topologies, the router will not be able to resolve a next hop.

The command for configuring a default static route is a global configuration command. The correct syntax was shown earlier in the “Static Routes” section, because it is configured in the same way as a static route. However, the source and destination prefix is set to 0.0.0.0 0.0.0.0. If the default route has no knowledge of the destination route because there is no entry in the routing table, it does not care where the datagram came from or where it is destined and therefore sends it to the address defined in the default route.

Floating Static Routes

Floating static routes are another mechanism for entering information manually into the routing table. They also overcome a limitation or constraint in the network design. A floating static route allows a backup route to lie dormant until the primary route dies. Then the backup route is activated and takes the place of the primary route until the network is mended. At that point, the backup route returns to sleep until the next time it is called upon. An example is a dialup line that serves as the backup to the Frame Relay link connecting a campus to the core of the network.

The complication of this solution is that the backup route is statically defined and is therefore considered more reliable than the route learned through a routing protocol. This means that the backup link would immediately become the primary path and be entered into the routing table.

To be sure that the backup route never overrides the original route, it is defined as a floating static route, allowing it to be the primary link only in the event of the primary link’s untimely death. The

static link is made less attractive than the primary link, despite its status as a static route, by changing the administrative distance to a higher value than that of the primary link.

On Demand Routing

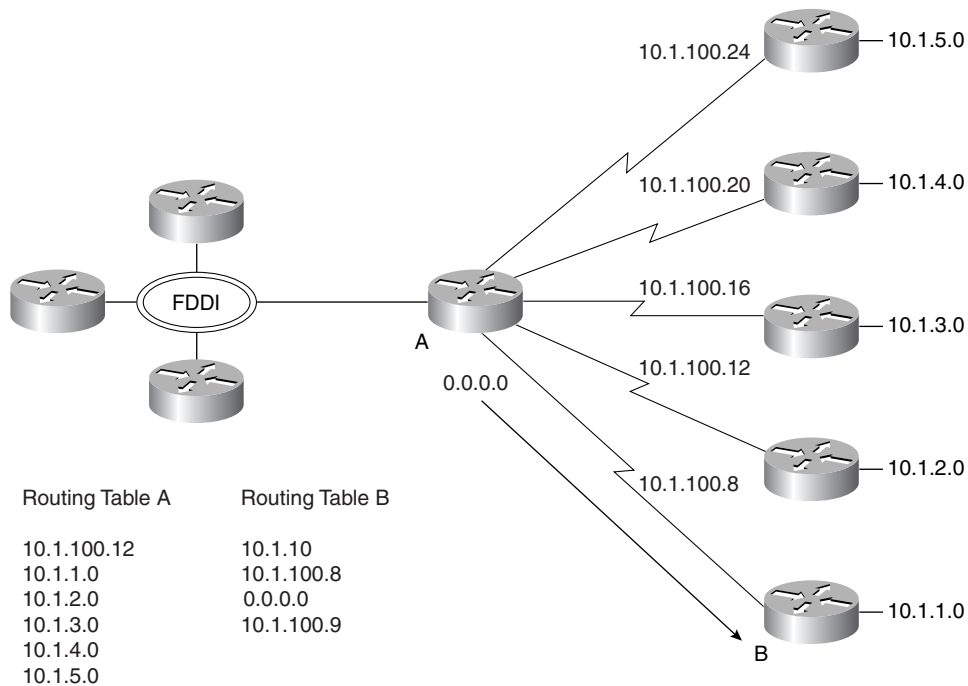
Everything comes at a cost. In the case of routing updates, static routing has a high administrative overhead, and dynamic routing consumes network resources. Normally, choosing whether to implement static or dynamic routing is a straightforward decision; high speed and cheap network resources reduce the concern nowadays. Static routing, while still used to define networks, is often used to share routing information between classful and classless routing protocols or to define a default route. However, in a network that has a large distributed topology, neither dynamic nor static might be appropriate. In such a network, the links are often low bandwidth and very little information might need to be sent across the link. It would appear that static and default routes are a prime candidate for this situation. However, if there are many remote sites in the typical hub and spoke configuration, this might become an unmanageable administrative burden. In On Demand Routing (ODR), all spoke routers can have an identical configuration, though obviously the IP addresses must be unique to each router.

ODR uses Cisco Discovery Protocol (CDP) to send the prefixes of attached networks from the spokes, or stub networks, to the hub or core router. The hub router sends its interface address of the shared link as the default route for the stub router.

ODR has the advantage of sending minimal information, such as the prefix and mask and the metric of one, every 60 seconds by default. This information populates the routing table of the hub router and can be redistributed into a routing protocol. Because the mask is sent in the update, VLSM can be used.

Figure 1-3 shows a typical example of when ODR could be used. In this figure, Router A, as the hub, has knowledge of all the networks connected to every spoke (stub router). The networks for the rest of the autonomous system have not been placed in Router A's routing table in order to simplify the figure.

Every spoke router, represented here by Router B, is sent a default route to the rest of the network. The default route is 0.0.0.0, with the next hop being the IP address of the interface that connects Router A to each spoke router. Router B, therefore, has two networks to which it is directly connected: the default network 0.0.0.0, and the next hop address to the default network, which is the IP address of Router A.

Figure 1-3 *Typical Network Topology for ODR*

When configuring ODR, it is important to remember the following points:

- No IP routing protocol should be configured on the stub routers; IP routing is on by default, allowing the use of default routes.
- Any secondary addresses configured on the stub routers are not propagated by CDP to the hub router.
- ODR must be configured on the hub router.
- Although CDP is enabled by default on interfaces, some WAN interfaces, such as ATM, require CDP to be configured on the interface with the command **cdp enable**.
- CDP uses multicast; therefore, on WAN technologies that require mapping statements, use the broadcast statement to ensure the CDP updates are propagated.
- It might be necessary to tune the CDP timers to send updates more frequently than every 60 seconds.

To configure CDP/ODR, use the commands described in Table 1-7.

Table 1-7 ODR Configuration Commands

Configuration Mode	Command	Purpose
Router(config-if)#	cdp timer <i>seconds</i>	Changes how often the CDP updates are sent
Router(config)#	router odr	Enables ODR on the interface and is only configured on the hub router
Router(config)#	timers basic update invalid holddown flush [<i>sleeptime</i>]	Changes how often the ODR routes are flushed from the routing table

Routing Versus Switching

This section describes the roles of routing and switching and how these functions work together to move datagrams through the network. Cisco makes a distinction between the routing function and the switching function of a router. The difference is simple: Two jobs within a router need to be done to move a datagram from an incoming interface to the outgoing interface. The destination path first must be determined, and then the datagram can be sent to the outgoing interface and on to the next stage in its journey. The path determination is the routing function, while sending the datagram to the outgoing interface is the job of the switching process.

The Routing Function

The *routing* function is responsible for learning the logical topology of the network and then making decisions based on that knowledge. The decisions determine whether the incoming datagram can be routed and, if so, how.

The routing process goes through these steps when a datagram is received. The steps to route a datagram can be summarized into the following questions:

1. Is the protocol stack and routing protocol configured on the router?
2. If the protocol stack is present, is there an entry for the remote network in the routing table?
3. If the destination network is not in the routing table, is there a default network configured?
4. If either a dynamic or a default route is available, is the destination network reachable?
5. Which is the best path to that remote network?
6. Are there multiple equal-cost paths?
7. If there are multiple equal-cost paths, to which outgoing interfaces should the datagrams be queued?

The Switching Function

The *switching* function is concerned with moving data across the router. It is responsible for forwarding the datagram. Switching takes over after the routing decisions have been made. Although the router has lookups to make, the few decisions that need to be made are performed in hardware. Therefore, this function is very fast.

The switching function does the following:

1. Checks the incoming frame for validity
2. Checks whether the frame is addressed (at Layer 2) to the router
3. Checks whether the frame is within the scope of the framing criteria (too big or too small)
4. Checks whether the frame passes cyclic redundancy check (CRC)
5. Strips the Layer 2 header and trailer from the frame and checks the destination address against the cache entries
6. Creates the appropriate frame header and trailer (if there is an entry in cache for the destination address) and forwards the frame to the outbound interface queue

NOTE Note that the preceding sections refer to the internals of the Cisco IOS software, which are extremely complex. The software has been described at the level required by the BSCI exam. If you are designing and configuring live networks, it would be wise to understand some of the issues in more depth. This information is readily available on the Cisco web site.

Functionality is broken into two components to ensure that the process is as fast as possible. After the routing decisions are made, the Cisco router caches the result, allowing subsequent datagrams to be switched.

The Routing/Switching Relationship in a Cisco Router

A datagram transiting the router is accepted into the router if the frame header (of the frame in which the datagram resides) contains the Layer 2 address of one of the router's interfaces. If properly addressed, after the framing is checked, the frame and its content (the datagram) are buffered, pending further processing. The buffering occurs in main memory or some other specialized memory location.

If the source and destination Layer 3 address of the datagram have not been seen by this router before, the datagram will be process switched, or routed. This involves the following actions:

1. When a datagram is to be forwarded, a process initiates a lookup in this routing table and a decision about how the datagram should be forwarded.

2. The datagram is then encapsulated.
3. If fast switching is enabled, the datagram is examined again, and an entry is put into a route cache. The entry in this cache consists of the following:
 - An IP prefix
 - The output interface
 - The link-layer header to be used in forwarding the datagram

On subsequent datagrams, if the IP destination matches a prefix found in the route cache, the datagram is forwarded using this information. The routing function is not disturbed, nor are the CPU cycles required to feed this monster expended.

The type of route cache used depends on the hardware used. The caches available are called *fast switching*, *autonomous switching*, *silicon switching*, and *Cisco Express Forwarding (CEF)*.

Foundation Summary

The “Foundation Summary” section of each chapter lists the most important facts from the chapter. Although this section does not list every fact from the chapter that will be on your exam, a well-prepared candidate should, at a minimum, know all the details in each “Foundation Summary” before going to take the exam.

Table 1-8 shows the metrics used by the IP routing protocols.

Table 1-8 *Routing Protocol Metrics*

Routing Protocol	Metric
RIPv1	Hop count
RIPv2	Hop count
IGRP	Bandwidth, delay, load, reliability
EIGRP	Bandwidth, delay, load, reliability
OSPF	Cost (The Cisco default states that the cost of an interface is inversely proportional to the bandwidth of that interface. A higher bandwidth indicates a lower cost.)
IS-IS	Cost

Table 1-9 explains how to read the information in the routing table, as explained in the **show ip route** command.

Table 1-9 *Explanation of the show ip route Command*

Code	Protocol That Derived the Route
I	IGRP
D	EIGRP
D EX	External EIGRP
R	RIP
C	Directly Connected
S	Static
E	EGP
B	BGP

Table 1-9 *Explanation of the show ip route Command (Continued)*

Code	Protocol That Derived the Route
i	IS-IS
I L1	IS-IS level 1
I L2	IS-IS level 2
M	Mobile
U	Per-user static route
o	ODR
T	Traffic-engineered route
O	OSPF networks from within the same area as the router. These are networks learned from router and network LSAs.
O IA	OSPF interarea. This is sent out by the ABRs and is created from the summary link LSA (type 3 and type 4). These routes will not be seen on a router within a totally stubby area because it will not receive LSAs external to the area.
O N1	OSPF NSSA external type 1
O N2	OSPF NSSA external type 2
O E1	OSPF external type 1. These routes are generated by the ASBR and show routes that are external to the autonomous system. The cost of this external route is the summarization of the external cost, plus the cost of the path to the ASBR. These routes will not be seen in a stub or totally stubby area.
O E2	OSPF external type 2. These routes do not take into account the cost of the path to the ASBR. They consider only the external cost.

Q&A

As mentioned in the introduction, “All About the CCNP, CCDP, and CCIP Certifications,” you have two choices for review questions. The questions that follow next give you a bigger challenge than the exam itself by using an open-ended question format. By reviewing now with this more difficult question format, you can exercise your memory better and prove your conceptual and factual knowledge of this chapter. The answers to these questions are found in Appendix A.

For more practice with examlike question formats, including questions using a router simulator and multichoice questions, use the exam engine on the CD-ROM.

1. In the routing table, a field indicates the source of the routing information. If the field showed the letter C, what would this mean?
2. In the routing table, how is the next hop indicated?
3. Cisco distinguishes between the routing and the switching functions. What is the difference?
4. Name the interior IP routing protocols that send the mask with the routing update.
5. Does VLSM require a classful or classless routing protocol, and why?
6. State one of the characteristics of a classful routing protocol.
7. What is the command to show whether a specific network, such as 141.131.6.16, is present in the routing table?
8. State one major difference between a classful protocol and a classless routing protocol.
9. Describe briefly the difference between a routing and routed protocol.
10. Describe the processes used to build and maintain the routing table.
11. Give a brief explanation of the switching function.
12. What information is contained in the fast switching cache?
13. When would you use the command **no auto-summary** ?
14. When does OSPF send updates, and do they contain the entire routing table?
15. Network convergence is when a network domain has learned about the new network topology after a change has occurred in the domain. What is considered a network change that would require network convergence?
16. When would you consider using a static route in your network?
17. When would you consider using a default route in your network?

18. What command is used to configure ODR on a hub router?
19. When would you consider using a floating static route in your network?
20. Describe some of the characteristics of classless routing protocols.

Scenarios

The following scenario and questions are designed to draw together the content of the chapter and to exercise your understanding of the concepts. There is not necessarily a right answer. The thought process and practice in manipulating the concepts is the goal of this section. The answers to the scenario questions are found at the end of this chapter.

Scenario 1-1

Scenario 1-1 provides you with the data produced with the Cisco router IOS command **show ip route**. A legend defining the fields of the sample output is also provided to assist in answering the questions for Scenario 1-1.

Example 1-2 contains sample output from the **show ip route** command.

Example 1-2 The **show ip route** Command Output

```
Router#show ip route
Codes: I - IGRP derived, R - RIP derived, H - Hello derived, O - OSPF derived
       C - connected, S - static, E - EGP derived, B - BGP derived
       * - candidate default route, IA - OSPF inter area route
       E1 - OSPF external type 1 route, E2 - OSPF external type 2 route

Gateway of last resort is 131.119.254.240 to network 129.140.0.0

O E2 150.150.0.0 [160/5] via 131.119.254.6, 0:01:00, Ethernet2
E      192.67.131.0 [200/128] via 131.119.254.244, 0:02:22, Ethernet2
O E2 192.68.132.0 [160/5] via 131.119.254.6, 0:00:59, Ethernet2
O E2 130.130.0.0 [160/5] via 131.119.254.6, 0:00:59, Ethernet2
E      128.128.0.0 [200/128] via 131.119.254.244, 0:02:22, Ethernet2
E      129.129.0.0 [200/129] via 131.119.254.240, 0:02:22, Ethernet2
E      192.65.129.0 [200/128] via 131.119.254.244, 0:02:22, Ethernet2
E      131.131.0.0 [200/128] via 131.119.254.244, 0:02:22, Ethernet2
E      192.75.139.0 [200/129] via 131.119.254.240, 0:02:23, Ethernet2
```

The following information defines the fields reported in the **show ip route** command:

- The first column lists the protocol that derived the route.
- The second column might list certain protocol-specific information as defined in the display header.
- The third column lists the address of the remote network. The first number in the brackets is the administrative distance of the information source; the second number is the metric for the route.

- The fourth column specifies the address of the router that can build a route to the specified remote network.
- The fifth column specifies the last time that the route was updated, in hours:minutes:seconds.
- The final column specifies the interface through which the specified network can be reached.

Answer the following questions by using the output from the preceding **show ip route** command.

1. What routing protocol derived the route 130.130.0.0?
2. What router interface IP address is used to reach IP network 192.67.131.0?
3. When was the last time that the route 192.65.129.0 was updated?
4. Through which router interface can the IP network 128.128.0.0 be reached?

Scenario Answers

The answers provided in this section are not necessarily the only possible answers to the questions. The questions are designed to test your knowledge and to give practical exercise in certain key areas. This section is intended to test and exercise skills and concepts detailed in the body of this chapter.

If your answer is different, ask yourself whether it follows the tenets explained in the answers provided. Your answer is correct not if it matches the solution provided in the book, but rather if it has included the principles of design laid out in the chapter.

In this way, the testing provided in these scenarios is deeper: It examines not only your knowledge, but also your understanding and ability to apply that knowledge to problems.

If you do not get the correct answer, refer back to the text and review the subject tested. Be certain also to review your notes on the question to ensure that you understand the principles of the subject.

Scenario 1-1 Answers

1. *What routing protocol derived the route 130.130.0.0?*

OSPF.

2. *What router interface IP address is used to reach IP network 192.67.131.0?*

131.119.254.244. The fourth column of the sample output specifies the address of the router that can build a route to the specified remote network.

3. *When was the last time that the route 192.65.129.0 was updated?*

0:02:22. The fifth column of the sample output specifies the last time the route was updated, in hours:minutes:seconds.

4. *Through which router interface can the IP network 128.128.0.0 be reached?*

Ethernet2. The last column in the sample output specifies the interface through which the specified network can be reached.



This chapter covers the following topics, which you need to understand to pass the CCNP/CCDP/CCIP BSCI exam:

- IP subnetting
- Prefix routing/CIDR
- Variable-length subnet masks
- Summarization

IP Addressing

This chapter reviews some of the fundamental concepts of IP addressing. After mastering its subtleties, you will be able to consider the network management power that results from careful design of the IP addressing scheme.

The discussion of IP addressing is fundamental to the rest of the topics covered within this book and will directly reflect questions on the exam. If you do not thoroughly understand the contents of this chapter, it is safe to say that it will be impossible to pass the exam. The entire BSCI exam is based on IP addressing. For this reason, IP addressing is reviewed thoroughly in this book.

“Do I Know This Already?” Quiz

The purpose of the “Do I Know This Already?” quiz is to help you decide what parts of this chapter to use. If you already intend to read the entire chapter, you do not necessarily need to answer these questions now.

The 12-question quiz, derived from the major sections in the “Foundation Topics” portion of the chapter, helps you to determine how to spend your limited study time.

Table 2-1 outlines the major topics discussed in this chapter and the “Do I Know This Already?” quiz questions that correspond to those topics.

Table 2-1 “Do I Know This Already?” Foundation Topics Section-to-Question Mapping

Foundation Topics Section	Questions Covered in This Section
IP Subnetting	1–3
Prefix routing/CIDR	4–6
Variable-Length Subnet Masks	7–9
Summarization	10–12

NOTE The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark this question wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

1. You are given a Class C address with the requirement to accommodate 14 subnets, not including the zero address for the network or the broadcast address for the subnet. It is also necessary to have 10 hosts on each subnet. What subnet mask would you use?
 - a. 255.255.255.240 with a prefix of /28
 - b. 255.255.255.248, which is a prefix of /28
 - c. 0.0.0.15, which is a prefix mask of /28
 - d. 255.255.255.0, which is the network number which is no prefix mask
2. List the range of hosts available on the 136.122.10.192/28 subnet. Do not include the zero address or the broadcast address for the subnet.
 - a. 136.122.10.193–136.122.10.223
 - b. 136.122.10.193–136.122.10.206
 - c. 136.122.10.192–136.122.10.254
 - d. 136.122.10.192–136.122.10.206
3. Convert the subnet address 56.98.5.0/24 to binary notation.
 - a. 00111010.01100110.00000101.00000000
 - b. 00111000.01100110.00000101.00000000
 - c. 00111000.01100010.00001001.00000000
 - d. 00111000.01100010.00000101.00000000
4. What does *CIDR* stand for?
 - a. Classful Interdomain Resolution
 - b. Classless Intradomain Routing
 - c. Classless Interdomain Routing
 - d. Classful Interdomain Routing

5. From the following list, which is an advantage of prefix routing?
 - a. The routing tables in the Internet are smaller.
 - b. The routing tables within an autonomous system are smaller.
 - c. The Internet has to allocate fewer Class C addresses.
 - d. Prefix routing adds to the IP header, which speeds up the fast cache forwarding process.
6. The address 202.100.48.0/21 can be identified as which of the following?
 - a. A Class C address
 - b. A group of eight Class C networks summarized to one address
 - c. A Class B address subnetted to 5 bits, allowing 30 subnets
 - d. A mistake in subnetting
7. Which routing protocols support VLSM?
 - a. RIPv2, OSPF, IS-IS, EIGRP, BGP-4
 - b. RIPv2, OSPF, IS-IS, IGRP, EIGRP, BGP-4
 - c. RIPv2, OSPF, IS-IS, EIGRP
 - d. RIPv2, OSPF, EIGRP, BGP-4
8. What does VLSM stand for?
 - a. Various-length subnetting mask
 - b. Variable-length subnet mask
 - c. Very long sustained mantra
 - d. Variable-length subnet mechanism
9. Which of the following options best describes the difference between CIDR and VLSM?
 - a. CIDR is the technique for designing and allocating addresses, whereas VLSM defines the technology used by the routing protocol.
 - b. VLSM summarizes networks within the Internet.
 - c. There is no difference between these terms; they are interchangeable.
 - d. CIDR summarizes networks within the Internet as it is applied to blocks of classful networks. VLSM is used within an autonomous system, as it is applied to a classful network.

10. The Class B network address of 133.222.0.0 has been given a mask of 255.255.255.0. The subnets 133.222.8.0, 133.222.9.0, 133.222.10.0, 133.222.11.0, 133.222.12.0, 133.222.13.0, 133.222.14.0, and 133.222.15.0 need to be summarized using VLSM. Give the subnet and new mask to achieve this summarization.
 - a. 133.222.16.0/21 with a mask 255.255.240.0
 - b. 133.222.16.0/20 with a mask 255.255.240.0
 - c. 133.222.8.0/21 with a mask 255.255.248.0
 - d. 133.222.8.0/20 with a mask 255.255.240.0
11. Given the mask 255.255.255.224 for the IP network 131.108.16.64, which of the following subnets are summarized within this address?
 - a. 131.108.16.92
 - b. 131.108.16.96
 - c. 131.108.16.65
 - d. 131.108.16.81
12. Manual summarization is available in which of the following interior routing protocols?
 - a. OSPF, RIPv1, IS-IS
 - b. EIGRP, OSPF, BGP-4
 - c. EIGRP, OSPF, RIPv2
 - d. EIGRP, IS-IS, RIPv2, OSPF

The answers to this quiz are found in Appendix A, “Answers to Chapter ‘Do I Know This Already?’ Quizzes and Q&A Sections.” The suggested choices for your next step are as follows:

- **6 or less overall score** —Read the entire chapter. This includes the “Foundation Topics” and “Foundation Summary” sections, the “Q&A” section, and the “Scenarios” at the end of the chapter.
- **7 to 9 overall score** —Begin with the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. If you have trouble with these exercises, read the appropriate sections in “Foundation Topics.”
- **10 or more overall score** —If you want more review on these topics, skip to the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. Otherwise, move to the next chapter.

Foundation Topics

IP Subnetting

IP subnetting is the means by which IP networks are addressed. It refers to the IP Layer 3 address, but it also refers to how one large IP network address is subdivided into many smaller network addresses. This section introduces the concept of addressing. It explains the need for a Layer 3 address in a routed network and the relationship between a Layer 3 and a Layer 2 network address. Against this background, the section explains IP addressing and how to address an entire organization with many networks from one unique network address.

The Need for Layer 3 Addressing

A Layer 3 address is a logical address sitting on top of a physical network structure. Its importance is in its ability to direct traffic and thus overcome the need for broadcasts, which can cause problems for a switched environment. A Layer 2 (switched) network cannot control or limit broadcasts, which can therefore saturate the available bandwidth. The resulting congestion results in slow response times and the loss of sessions.

A Layer 3 address allows network traffic to be directed to its destination. In fact, the purpose of any address is to find a specific location, whether it is the address of a restaurant or the company e-mail server. The location is found because every address is hierarchical; just as a restaurant is in a city, on a street, at a street number, the e-mail server is on a network at a host number.

Network Structures and Data Flow

For data to be sent to its destination, the underlying physical structure, or wiring, should support the logical structure, or the Layer 3 addressing. This structure also should reflect the organizational data flow. It would make sense for servers to be accessible to departments that share information and for the physical wiring and logical addressing to support this sharing of resources. Therefore, the servers might be physically adjacent and on the same IP subnet. Both the physical and the logical structure of the network should support the organizational data flow because without this structure, application data can wander throughout your network inefficiently, clogging up available bandwidth.

The Network and How It Is Addressed

Layer 3 provides the capability to logically address the network. To appreciate fully the power and purpose of the Layer 3 address, it is important to understand the meaning of the term *network* (as defined by Layer 3).

A network address has two parts: the network and the host portions. The host portion of the address identifies the individual device within a group. The network portion identifies a group of individual devices.

Unfortunately, the term *network* is used loosely; although it is often defined, the term is seldom understood. In addition, the term *network* appears in several different contexts, compounding the confusion created for the user.

It is increasingly important to have an accurate definition of a network because new technologies, such as VLANs and Layer 3 switching, have blurred the distinctions between the different layers of the OSI model.

Layer 3 switching and VLANs are a technology that use an intelligent switch to distinguish between different logical networks at Layer 3. It is possible, therefore, to transfer data at great speed in hardware, because no routing decisions need to be made. If data needs to be transferred between logical networks, a routing decision at Layer 3 needs to be made, which will take longer.

The following list outlines the various uses of the term *network*:

- The piece of wire or physical medium to which a group of devices are connected. This is more accurately defined as a segment.
- A Layer 3 network.
- The LAN.
- The corporate or organizational network.

For our purposes, the term *network* refers to the Layer 3 network.

Layer 3 Network Characteristics

A Layer 3 address is a logical address imposed on a physical network with physical addresses hardcoded into the devices. The logical address is one that is created by the administrator to allow data to be directed through the network to the remote destination. A Layer 3 address comes in two parts, the network and the host. The network portion of a Layer 3 address is a border chosen by an administrator to group end devices. This group is given an identifier or label, which is the network number.

A Layer 3 network address has the following characteristics:

- The network number defines a group of end devices or hosts and labels the group with a network number.
- The address is hierarchical, which allows decisions to be made on groups of devices.

- The devices running the Layer 3 protocol do not forward broadcasts.
- The group address combined with the unique membership number for that group identifies the end device. This is the host address.
- Although the identifier for the end device might not be unique to the organization, it will be unique to the group or network.
- If the addressing is carefully planned and the addressing scheme allows, groups can be consolidated (cities into states, states into countries, countries into continents, for example). The networks logically grouped together under one administrative control, such as a company, are called an autonomous system.

An Analogy for Understanding the Concept of a Network

Administrative lines, or borders, are drawn between one city and another, between one state and another, and even between countries. These borders serve the same purpose as the network portion of a Layer 3 address; that is, they allow rules to be placed on a group of end systems (in the geographic analogy, humans).

With a logical Layer 3 address, the network can direct traffic to specific devices. Routing tables, which are lists of networks held in routers, serve as maps and road signs.

It is very important to plan carefully the placement of these boundaries to ensure the geographic proximity of the end devices or hosts. After boundaries are defined, they seldom change. This is not to say that they *cannot* change; indeed, historically, boundaries between cities, states, and countries have been redefined, but not without careful thought and the possibility of some transitional trauma. With the emergence of VLANs, however, it is easier to change a network boundary.

Layer 3 to Layer 2 Conversion

Although it is important to understand the need for a Layer 3 address, it becomes much easier to understand in the context of host-to-host communication. This requires a brief journey back to the OSI seven-layer model.

When an end system or host decides to send data to another system, certain things have to happen. The application generates the data and hands it down the stack until the Layer 3 address and packet header is added. The appropriate Layer 2 adds the header and Layer 2 address, which is an address with no hierarchy; that is, it is a flat address. No hierarchy is needed because the destination machine is either directly connected to the same medium or on the same technology, such as Frame Relay. The Layer 2 sends the frame to the physical layer, where it is transmitted to the destination end device.

On receiving the incoming bits, the destination system will buffer the bits until a frame is assembled. Layer 2 will ask the following questions and perform the following tasks:

- Is the frame valid?
- Does it pass the cyclic redundancy check (CRC)?
- Is it too small or too big?
- Is the frame addressed to this device (at Layer 2)?
- For which Layer 3 protocol is the frame destined (for example, IP and IPX)?
- Is that Layer 3 running on the device?
- Strip off the Layer 2 header and address.
- Pass the packet or datagram up to Layer 3.

At this point, Layer 3 will ask the following questions and perform the following tasks:

- Is this datagram or packet addressed to me?
- If the packet is addressed to me, and if it has not been damaged in transit but passes the Layer 3 verification, then strip off the header and pass it up to the upper layer.
- If the frame is not valid, then drop the frame.
- If the packet is not addressed to this system and the system is a router, the packet is handed to the switch or routing process.
- The router will first look in all the caches to ascertain whether packets have been forwarded to this Layer 3 address before. If this is the case, the address is likely to be cached, allowing the datagram to be switched to its next hop. If the address is not in cache, the packet is sent to the routing process to be process switched. Subsequent packets will be switched.

An IP Address

TCP/IP is unique because, although it has a fixed 32-bit address, it does not have a fixed number of bits allocated to the network or host portion of the address, in the way that AppleTalk or IPX addresses were defined. Therefore, an IP address can be read only in the context of a subnet mask.

A governing body, the *Internet Assigned Numbers Authority* (IANA, <http://www.iana.org/>), allocates an original address. This address can be subdivided into a range of networks called subnets by reallocating the host bits as network bits. The number of bits reassigned to be network bits is dependent on the number of networks that are required. To identify how many of the address bits have been extended into the network portion of the address, a subnet mask is used.

The subnet mask defines the network portion of the address, by masking or obscuring the host portion of the address, revealing the network address. The subnet mask is therefore crucial to the

ability to route traffic, as every router needs to identify the network portion of the destination address in order to forward the packet.

Unfortunately, IP network terminology is vague, and the address provided by the Internet community might be referred to by any of the following terms:

- Address provided by the IANA
- Classful address
- Supernet address
- Internet address
- Network address
- Major address

NOTE For the purposes of this book, the term *classful address* is used to refer to the unique network address given out by the IANA body. It might also be referred to as the *IANA address*.

The Internet community originally identified three classes of organizations:

- Small organizations fall into Class C
- Medium organizations fall into Class B
- Large organizations fall into Class A

Actually, five classes of addresses are used on the Internet. The other two classes represent multicast (Class D) and experimental addresses (Class E). Routing protocols and videoconferencing increasingly use Class D addresses.

A router identifies the class of address by looking at the first few bits of the 32-bit address. When looking at the address in a decimal format, the number in the first octet reveals the class of address. This is known as the first octet rule.

Table 2-2 shows how the classes are broken up.

Table 2-2 *Classes of Addresses*

Class of Address	First Octet	Number of Hosts That Address Could Represent on One Network
Class A address	001 to 127	Could represent 16.77 million hosts on one network; the 127 address is reserved as a loopback address
Class B address	128 to 191	Could represent 65,534 hosts on one network

(continues)

Table 2-2 *Classes of Addresses (Continued)*

Class of Address	First Octet	Number of Hosts That Address Could Represent on One Network
Class C address	192 to 223	Could represent 254 hosts on one network
Class D address	224 to 239	Not relevant
Class E address	240 to 254	Not relevant

The Internet Authoritative Bodies

The Internet community assigns an organization with a unique binary pattern or classful address. The group within the Internet community responsible for allocating unique classful networks has changed over the years. Originally, the government-funded IANA assigned numbers and was, until recently, commercially administered by Networks Solutions of Herndon, Virginia. On November 25, 1998, the Internet Corporation for Assigned Names and Numbers (ICANN) was officially recognized. This global nonprofit corporation, currently managed by the U.S. government, was created to perform administrative functions for the Internet. ICANN has gradually taken over responsibility for coordinating the assignment of protocol parameters, the management of the domain name and root server systems, and the allocation of IP address space.

The growth of the Internet has led to regional organizations for the allocation of IP addresses, and under ICANN, the IANA continues to distribute addresses to the regional Internet registries.

The most recent list of these follows:

- Regional registries:
 - Asia-Pacific Network Information Center (APNIC), <http://www.apnic.net>
 - American Registry for Internet Numbers (ARIN), <http://www.arin.net>
 - Réseaux IP Européens (RIPE), <http://www.ripe.net>
- Domain registration: InterNIC, <http://www.internic.net>

When it has possession of the classful address, an organization is responsible for determining where to place the boundary between network and host addresses and is responsible for addressing the network.

Allocating the network address bits is a straightforward task because it is simply a matter of counting—counting bits and counting in binary, but counting nonetheless. In addition, many charts can help ease the pain of binary-to-decimal translation. Although it is easy to implement, the complexity lies in the network design.

The following example illustrates how the bit allocation of a subnet address works.

An Example of Bit Allocation in a Network Address

If 10 bits are allocated to the network portion of the address, 22 bits are left to the host portion of the address. In binary, 10 bits can be used to represent 1024 distinct entities or networks (each being assigned a unique bit pattern or address). The 22 bits left to identify hosts can be used to represent four million hosts (actually, 4,194,304) on each network.

The total number of devices that can be addressed is calculated by multiplying the number of host addresses available on each network by the number of networks that can be addressed, as follows:

$$4,194,304 * 1,024 = 4,294,967,296$$

However, the administrator does not have the whole 32 bits to use. The Internet community, which manages the addresses to ensure their uniqueness, allocates a unique bit pattern to each organization that requests a connection to the Internet. This bit pattern is then used to uniquely identify the organization within the Internet.

The Subnet Mask

The way that the *subnet mask* extracts the network portion of the address from the whole IP address is by using a logical AND operation. Once you understand the principles of how this works, the math is easy.

The Logical AND

When an address is assigned to an interface, it is configured with the subnet mask. Although represented in a *dotted decimal* form, the router converts the address and mask into binary and performs a logical AND operation to find the network portion of the address.

To perform a *logical AND*, the IP address is written out in binary, with the subnet or Internet mask written beneath it in binary. Each binary digit of the address is then ANDed with the corresponding binary digit of the mask.

The rules of the AND operation are as follows:

- Positive AND positive is positive.
- Negative AND anything is negative.

This means that the following is true:

- 1 AND 1 is 1.
- 1 AND 0 is 0.
- 0 AND 1 is 0.
- 0 AND 0 is 0.

Figure 2-1 illustrates the AND logic.

Figure 2-1 *AND Logic and the Subnetwork*

IP address	144.100.16.8	
IP subnet mask	255.255.255.0	
IP address in binary	10010000.01100100.00010000.	00001000
IP subnet mask in binary	11111111.11111111.11111111.	00000000
The result of the logical AND		10010000.01100100.00010000.
		00000000

Layer 3 can now make a decision on how to route the network number that has been revealed. The result is the removal of the host portion of the address, and the subnet address is left intact. Therefore, the host 144.100.16.8 is a member of the subnet 144.100.16.0, which is the result of the logical AND converted to decimal.

With this information, the router can now perform a search on the routing table to see whether it can route to the remote network. The routing table uses the network entry that has the longest match of bits to the destination network.

NOTE The terms used to describe the mask are numerous and often vague. This book uses the term *subnet mask* when referring to the mask used within an organization, and it uses *Internet mask* or *prefix mask* when referring to the address allocated by ARIN.

When determining the subnet mask, certain rules must be followed. RFC 950, “Internet Standard Subnetting Procedure,” outlines these rules.

NOTE You can find all RFCs online at <http://www.isi.edu/in-notes/rfcxxx.txt>, where *xxx* is the number of the RFC. If you do not know the number of the RFC, you can find it by doing a topic search at <http://www.rfc-editor.org/rfcsearch.html>.

Familiar Rules in IP Subnetting

Because originally the routing protocols could not send the subnet mask with the routing update, the first set of rules about applying IP addresses were different than they are now. For the most part,

these rules still hold true. With the advent of new technology, however, it is now possible to surmount some of the previous limitations set out in RFC 950.

The earlier (and perhaps familiar) rules included the following:

- The network bits (bits set to 1) in the subnet mask do not need to be contiguous, although they are advised to be contiguous.
- The network bits must not be all 0s or 1s.
- The decision on the number of bits allocated to the network is made once per classful address.

Because the original routing protocols did not send the subnet mask with the routing update, each router that received a subnet entry had to make some assumptions. The router assumed that the mask in use for the received subnet was the same as the one configured on its system.

If the subnet received in the routing update was of a different classful address (it was not configured on one of the router's interfaces), the router resolved the network address to the class address. The class of network was determined by the *first octet rule*. This rule uses the first few bits of the address to identify the class of address. The first octet rule is explained in more detail in the next section, "The New Subnet Rules."

When designing an IP network, you must ensure that the subnet mask is consistent in a classful network. As in a classful network, the routing protocol does not send the subnet mask in with the routing updates. If the subnet mask is not consistent, the routers might become confused and the network discontinuous.

New technology means that routing protocols can now send the subnet mask with the routing update. Therefore, the earlier rules regarding network classes do not necessarily apply.

The New Subnet Rules

Because the newer routing protocols can send the mask with the routing update, it is possible to have greater flexibility in the IP addressing design of your network. In particular, it is no longer necessary to adhere to the rule that the subnet mask can be created only once per classful network. The mask is held with the subnet in the routing table, which allows the distinction between the broadcast address and the subnet address that has been defined. This requires variable-length subnet masks (VLSM), which are described in the section titled "Variable-Length Subnet Masks." Likewise, it is no longer necessary for either the classful address or the individual organization to conform to the rules of classful routing. *Classful routing* occurs when the Layer 3 device observes the Internet address class boundaries of A, B, C, D, and E. It does this by using the first octet rule, as shown in the following table.

Table 2-3 *The First Octet Rule*

Bit Pattern	Class of Address	First Octet Range
0	A	0 to 127
10	B	128 to 191
110	C	192 to 223
1110	D	224 to 239
1111	E	240 to 255

A classful routing protocol does not transmit any information about the prefix length. It uses the first octet rule to determine the class of address, which is why the routing protocol cannot support VLSM. If the routing protocol is not connected to a classful network, it does not have a subnet mask, and it summarizes the address at the classful network boundary by using the first octet rule. Examples of classful routing protocols are Routing Information Protocol (RIPv1) and Interior Gateway Routing Protocol (IGRP).

The need for a subnet mask to summarize on a bit boundary other than the default provided by the first octet rule also prevents the summarization of class addresses within the Internet. However, if the routing protocol supports classless routing, there is no reason why Internet addresses cannot be summarized in the same way as subnets. As long as the address is allocated with a prefix mask to identify the network portion of the address, the IANA can hand out an address without regard for the bit boundary at Class A, B, or C.

The address must be allocated with a prefix mask to identify the network portion of the address. RFC 1812, “Requirements for IP Version 4 Routers,” restricts the flexibility of the addressing slightly, however, by requiring contiguous bits to be used in the mask.

It is also possible to overcome some of the rules regarding the allocation of network and host bits, which is explained later in the chapter in the section “Rules for VLSM.”

Prefix Routing/CIDR

Prefix routing, commonly known as classless interdomain routing (CIDR), is possible because of the newer routing protocols sending the subnet mask with the routing updates.

In this section, the need for CIDR and prefix routing is explained in the context of the problems experienced in the Internet with the size of routing tables. An example of how CIDR works is provided.

A Definition of Prefix Routing/CIDR

Prefix routing is just the means by which the Internet identifies the portion of the 32-bit TCP/IP address that uniquely identifies the organization. In effect, this means that the Internet can allocate a group of classful networks, represented by a single network address. This allows for prefix routing and summarization within the routing tables of the Internet. Prefix masks represent a group of TCP/IP network addresses using the method of address or subnet masks.

This aggregation of classful networks defies the old structure of Class A, B, C addressing, or classful addressing. The aggregation of classful networks, therefore, is classless and deals with connectivity between organizations through the Internet, referred to as *interdomain routing*. This technology is called CIDR. Table 2-4 shows the RFCs that outline the use of CIDR in an IP network.

Table 2-4 *RFCs about CIDR*

RFC Number	Title
1517	Applicability Statement for the Implementation of Classless Inter-Domain Routing (CIDR)
1518	An Architecture for IP Address Allocation with CIDR
1519	Classless Interdomain Routing (CIDR): An Address Assignment and Aggregation Strategy
1520	Exchanging Routing Information Across Provider Boundaries in the CIDR Environment

Problems with IP Addressing and the Internet

The Internet community found that small companies that wanted to connect to the Internet with a small number of hosts (50, for example) needed a Class C address, although a Class C designation might waste a large portion of its 254 addresses.

Conversely, if an organization has more than 254 hosts but fewer than 65,534 hosts, the Internet must either waste a large number of addresses by allocating a Class B address or provide multiple Class C addresses. RFC 1466, “Guidelines for Management of IP Address Space,” discusses the low percentage of allocated addresses in use.

The Class A, B, C address structure does not have enough granularity for today’s Internet. Because the Internet has grown in popularity, this has become a pressing problem. In addition, the number of entries in the routing tables of the Internet was reaching capacity, although only a small percentage of the addresses allocated were being used. The Internet started to reclaim unused addresses, but this was obviously a short-term solution. The implementation of CIDR with prefix routing is solving both problems, as you will learn about in the next sections.

CIDR as a Solution

An organization requiring multiple Class C addresses is allocated consecutive Class C addresses. However, the organization is issued only one address (representing the multiple addresses) for the Internet routing entry. This is achieved by pulling the network mask to the left, creating a prefix mask.

The shorter the prefix, the more generally the network is defined; the longer the prefix, the more specific the identification is. Table 2-5 visually demonstrates the use of the prefix.

Table 2-5 Table to Illustrate the Use of Prefix Masks

Prefix	Mask	New Address Space
/27	255.255.255.224	12 percent of a Class C 30 hosts
/26	255.255.255.192	24 percent of a Class C 62 hosts
/25	255.255.255.128	50 percent of a Class C 126 hosts
/23	255.255.254.0	2 Class Cs 510 hosts
/22	255.255.252.0	4 Class Cs 1022 hosts
/21	255.255.248.0	8 Class Cs 2046 hosts
/20	255.255.240.0	16 Class Cs 4094 hosts

The Internet IP addressing group ARIN, at <http://www.arin.net>, typically gives blocks of consecutive addresses to an Internet service provider (ISP) to allocate addresses to organizations that want to connect to the Internet. This reduces the routing tables even further by placing some of the address management responsibilities on the ISP.

CAUTION Connecting to an ISP requires some consideration because the ISP provides the addresses used in your organization. If you change your ISP, that address space will have to be relinquished back to the issuing ISP. This requires readdressing of the local network or some software application to translate the addresses. The Network Address Translation (NAT) is offered by Cisco and is an example of one such application, though there are many different solutions on the market.

In summary, CIDR solves the problem of the excessive network resources required to manage the huge routing tables. The next section provides an example of the use of CIDR.

An Example of the Use of CIDR

It is easy to see how CIDR works when the address and the mask are written in binary, as the router processes them. The Internet community has allocated a group of Class C addresses, although they are presented as a single network. Table 2-6 shows an example of an IP address in both decimal and binary format.

Table 2-6 *An IP Address and Mask Shown in Binary*

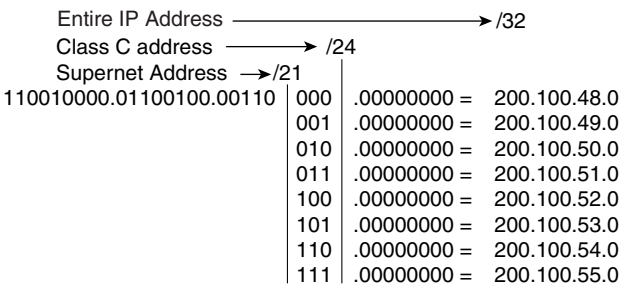
Description	Octet 1	Octet 2	Octet 3	Octet 4
IANA address in decimal	200	100	48	0
IANA address in binary	11001000	01100100	00110000	00000000
Prefix as a subnet mask in decimal	255	255	248	0
Prefix as a subnet mask in binary	11111111	11111111	11111000	00000000

If it were a standard Class C address, the mask would be 255.255.255.0. By making the mask 255.255.248.0, the last three bits of the third octet essentially give the organization eight Class C networks.

Imagine that a company called CyberKit has applied for a Class C address from the Internet authorities, though the company really needs a larger address space to address its network fully. To everyone's surprise, the company has been awarded eight Class C networks. The company owners are delighted because they were expecting only one Class C address.

Figure 2-2 shows the addresses awarded to CyberKit, the use of CIDR addresses, and how prefix routing works at the binary level.

Figure 2-2 Prefix Routing and the Use of CIDR



Although eight Class C addresses are provided to the organization, they are identified to the Internet as one address: 200.100.48.0, with a prefix mask of /21, which is the subnet mask of 255.255.248.0.

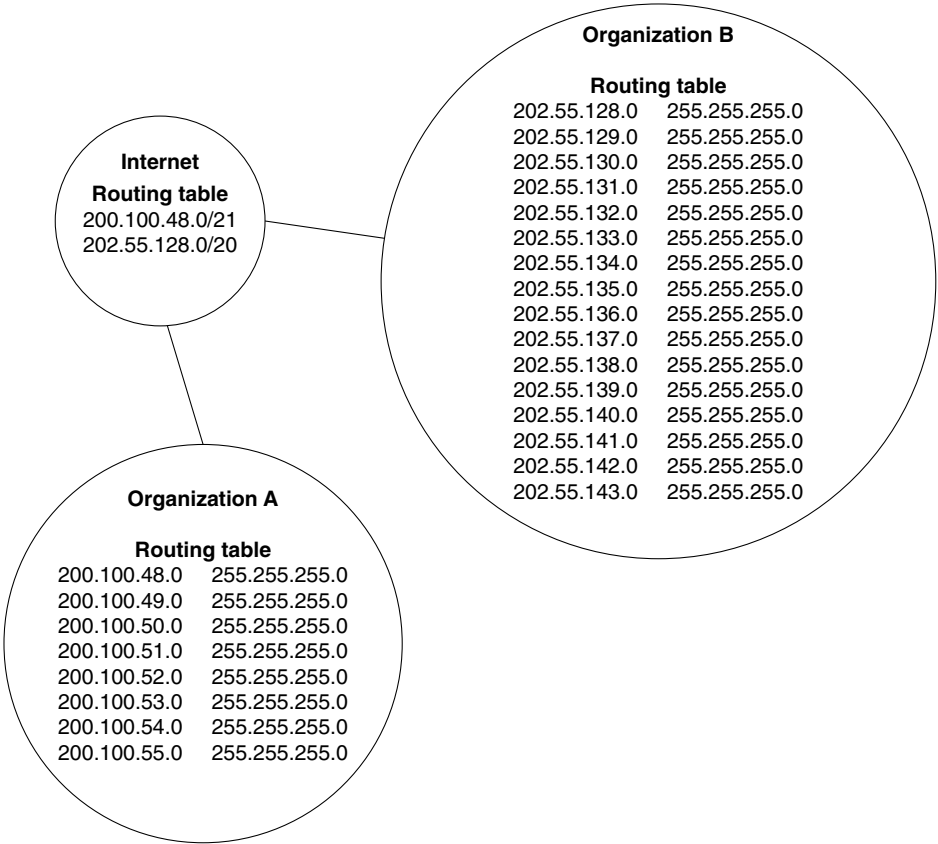
The organization does not have to use the addresses as Class C addresses. In accordance with the original rules, the organization can use the right-most zeroed bits however it deems appropriate.

Advantages of Prefix Routing/CIDR

Prefix routing is used to reduce the size of Internet routing tables. As explained in the preceding example, the Internet gave away the equivalent of eight Class C networks, but just one network entry appeared in the Internet’s routing table. In an environment that has more than 120,000 entries in the routing table (at the time of this writing), the size of the routing table in many ISPs has peaked at 120,000 entries. This is a significant reduction in the size of the routing table (which is expressed in terms of CPU utilization, memory, and bandwidth congestion).

In addition to the advantages of the original rules of TCP/IP addressing and subnet design, there is new flexibility granted to the Internet with prefix routing. The Internet no longer needs to abide by the rules of Classes A, B, and C. As shown, with some thought, many Internet networks might be presented as one network, thus reducing the network overhead. It could be said that the Internet has summarized many networks into one network. Figure 2-3 shows the effect of using prefix routing. The Internet’s routing table shows only two entries, 200.100.48.0 from organization A and 202.55.128.0 from organization B. This shows how the routing table within the Internet can be summarized, thus conserving resources.

Figure 2-3 Summarization of Internet Networks Using Prefix Routing



Prefix routing/CIDR or summarization achieves the same benefits in terms of the following:

- Reduction in the size of the routing table
- Less overhead in terms of network traffic, CPU, and memory
- Greater flexibility in addressing the networks

An organization can use summarization for the same reason as the Internet uses it with prefix routing: to reduce network overhead. The length of the prefix in this case depends on the number of bits needed rather than the Class A, B, and C structure.

NOTE The bit pattern provided by the Internet Assigned Numbers Authority (IANA) or any of its four Regional Internet Registries (RIRs) cannot be altered. The bits to the right of the unique address given by the IANA governing body are at the disposal of the organization.

To use the power of summarization within an organization, a sophisticated routing protocol that sends the mask with the routing updates is required. The capability to move the network/host boundary is called *VLSM*, which you will learn more about in the next section.

Variable-Length Subnet Masks

Variable-length subnet mask (VLSM) is used within an organization instead of CIDR, which is used within the Internet. VLSMs enable you to allocate required host bits on a granular basis.

Because organizations are rarely uniform in the distribution of hosts, it is much more efficient to provide only those host bits needed to address the number of hosts on a particular network.

An Example of VLSM

Consider a company that has been given a Class B address. The company has grown and now has some satellite offices that connect via point-to-point serial lines. The remote offices have eight workstations, three printers, and a router connecting them to the outside world. The main site has a building with ten floors, and each floor has approximately 25 workstations and four printers. A server farm in the basement has three servers and two routers. In this scenario, it is impossible to create a mask that serves all these environments. If you use an older routing protocol, you will waste a considerable amount of the available address space.

VLSM requires a routing protocol that supports the sending of the subnet mask.

The following routing protocols support VLSM:

- RIPv2
- OSPF
- IS-IS
- EIGRP
- BGP-4

Static routes could be said to use VLSM. They are often used when redistributing between routing protocols sharing an classful network when one routing protocol supports VLSM and the other does not. In these instances, the static route will define one summarized route for the non-VLSM routing protocol. This technique is also used when redistributing into BGP-4.

The following routing protocols do *not* support VLSM:

- RIPv1
- IGRP
- EGP

Rules for VLSM

The rules for variably subnetting an IP network are remarkably straightforward. The key is to remember that a hierarchical design in the addressing scheme is the goal. The physical network design also must reflect this logical hierarchy (Chapter 3, “Designing IP Networks,” discusses hierarchical design in detail). After the physical design is mapped, the logical structure can be placed on top of it.

The following rules apply when subnetting:

- A subnet can be used to address hosts, or it can be used for further subnetting.
- All 1s or all 0s in the subnet portion of the classful network could not originally be used, subsequently the command **ip subnet-zero** was introduced. It is a default setting for some Cisco equipment. If this rule has been followed, any subnet that is further subnetted does not need to obey this rule because it has already been observed.
- The routing protocol must carry the subnet mask in its updates.
- Multiple IP subnets intended for summarization must have the same high-order bits.
- Routing decisions are made on the entire subnet, and the router goes from more specific to more general when making routing decisions.

The two main reasons for using VLSM are that it makes efficient use of the available addressing and it enforces a good hierarchical design, allowing summarization and documentation.

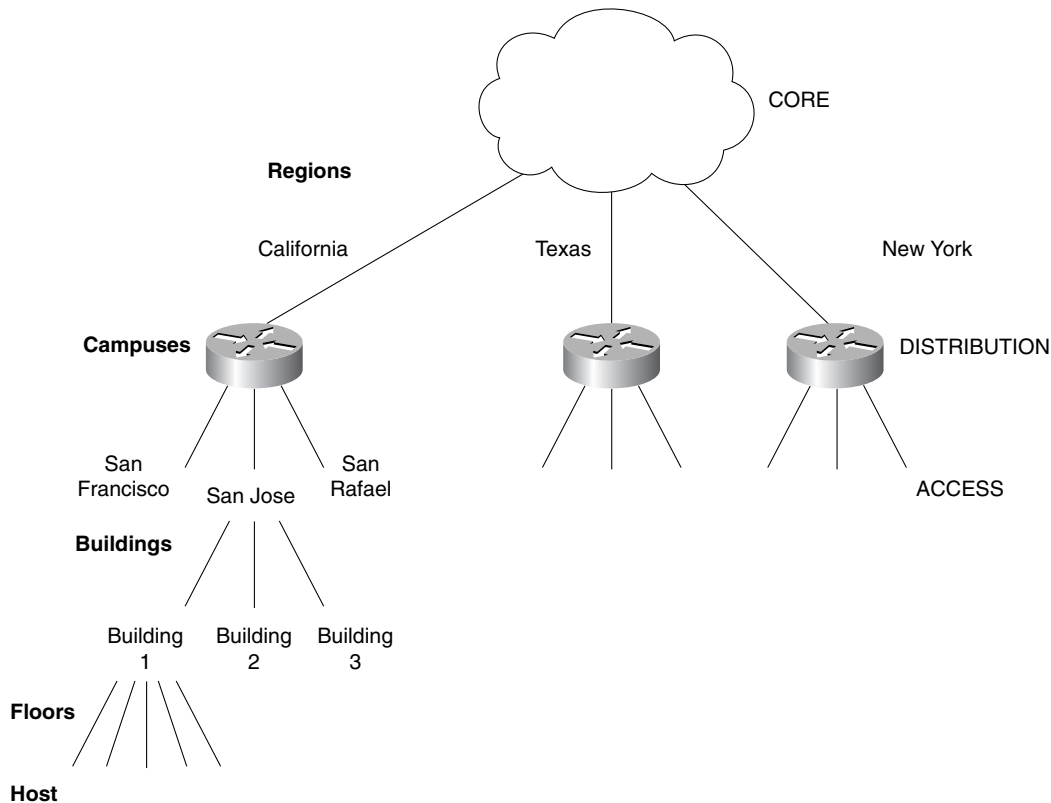
The benefits and the mechanics of VLSM are demonstrated in the following case study.

Case Study: Addressing the Network

To illustrate how VLSM works in supporting a hierarchical design and allowing summarization, this case study will break down a possible addressing scheme for a large organization.

To reassure you that it is actually a relatively easy task, a complicated example has been chosen. We will use a Class B address and create an addressing scheme for the company CyberKit.

If the Internet assigns the address 140.100.0.0, how might you address the network shown in Figure 2-4?

Figure 2-4 *Hierarchical Design of a Network Topology Used to Support the Use of VLSM*

The first task is to determine the number of regions, campuses, buildings, floors, and hosts on each floor. You also need to consider any anticipated growth or change in the network.

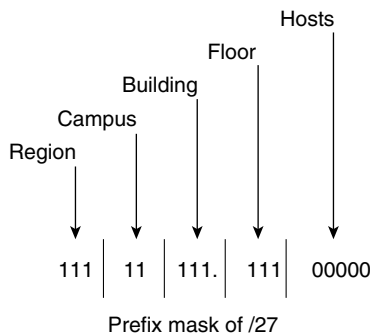
For this example, the network is comprised of the following:

- Three regions exist, but the company has plans to expand into other areas. Any expansion will probably not exceed eight states (adequate to cover the country).
- Within each region/state, there are no more than three campuses.
- Within each campus, there are no more than four buildings. This number might increase, however.
- No building has more than three floors.
- No floor has more than 30 hosts.

With this topology and growth detailed, it is possible to start allocating bits of the network address.

Taking the address 140.100.0.0 and writing out the last 16 bits, you can easily assign them to the different addressing tasks at hand. Figure 2-5 covers assigning IP addressing bits for VLSM.

Figure 2-5 *Assigning IP Addressing Bits for VLSM*



Consideration must be given to the subnetting rules (RFC 950 and RFC 1878, “Variable-Length Subnet Table For IPv4”) that state that there must not be all 0s or all 1s in the following:

- The Internet portion of the address
- The host portion of the address
- The algorithm for calculating the number of networks or hosts available is $2^n - 2$ (where n is the number of bits).

The subnet portion of the address used to be governed by this rule as well, but current Cisco technology allows the use of the all zero address for the subnet. The number of subnets is now calculated by the 2^n formula, where n is the number of bits by which the subnet mask was extended.

Historically, subnet zero was used by some network devices as a zero broadcast. Even today, some systems, such as Sun Solaris 4.x, have problems using subnet zero even with OSPF.

The command to enable the use of the 2^n zero subnet became the default configuration in version 12.0 of the Cisco IOS software.

However, you must still give attention to the host portion of the address. The host portion of the address must conform to the rule as defined; otherwise, it is not possible for the router to distinguish between hosts and broadcast addresses. An IP address cannot use all 0s or all 1s in the host portion of the address, because the all-0s address is used to show the subnet delimiter, and the all-1s address to broadcast to every device on the segment.

Allocating VLSM Addresses

Applying the addressing scheme designed in the preceding case study is very simple after the design has been worked out.

Taking California as the example to examine, we shall now address the entire region.

Figure 2-6 shows the bit allocation that was determined.

Figure 2-6 *Bit Allocation*

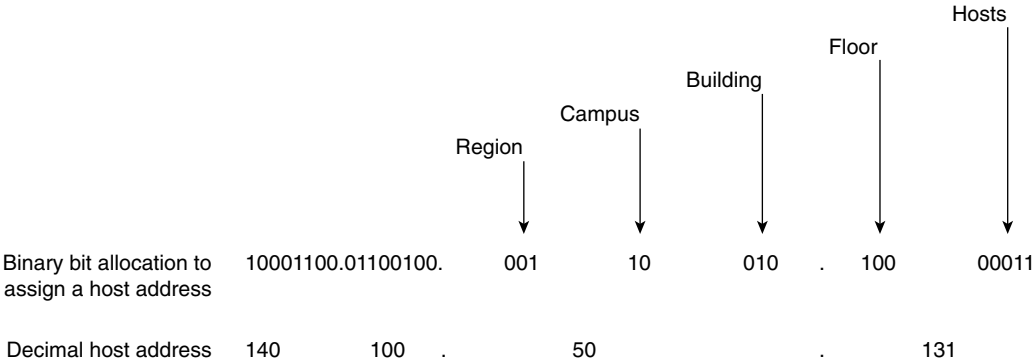
Region:
California: 001
Campus:
San Francisco: 01
San Jose: 10
San Rafael: 11
Buildings:
Building 1: 001
Building 2: 010
Building 3: 011
Building 4: 100
Floor:
Floor 1: 001
Floor 2: 010
Floor 3: 011
Floor 4: 100
Floor 5: 101
Hosts:
1-30

NOTE Remember that the case study will conform to the rule of reserving the broadcast addresses in the access layer of the network, the last level of subnetting.

Also remember that the buildings have the same bit pattern for each campus. However, this bit pattern is unique within the whole address space, because the pattern for the campus is unique and the address must be seen in its entirety.

The third host on the fourth floor of the second building in San Jose, California, will be given the address shown in Figure 2-7. The address in Figure 2-7 is represented as 140.100.50.131 in dotted decimal, with a mask of 255.255.255.224.

Figure 2-7 Example of How to Apply VLSM



Applying an addressing structure that uses VLSM with careful reference to the physical topology is very straightforward. When presented with a host address, it is common for people to try to determine the bit allocation working from the host address. If the addressing scheme has been well documented, network management is much easier, because as soon as the address is seen, its physical location is known. This simplifies troubleshooting, because a problem seen on a management console can be solved by member of the support staff.

NOTE This use of VLSM shows clearly that when allocating addresses in IP, it is necessary to reduce the address to binary and to disregard the octet boundary. Reducing the address to binary and disregarding the octet boundary creates a continuous set of bits to be applied as appropriate to address the network.

VLSM also enables you to allocate the required bits for addressing a particular network.

Optimizing the IP Address Space

Particularly in the use of WANs, where there is a predominance of point-to-point connections, allocating an entire subnet is very wasteful. VLSM allows refinement of the address space to exactly that which is needed and no more.

As demonstrated, dealing with VLSM to support the hierarchical design requires the consideration of the entire network topology. When using VLSM to optimize the IP address space, the network addressing can become extremely confusing if it is not clearly managed and documented.

In the preceding example, no consideration was given to the connections between the regions, campuses, and buildings—all of which could be point-to-point lines.

Now it is important to consider the last part of the network addressing, which will illustrate the use of VLSM for IP address optimization.

Assigning IP VLSM Subnets for WAN Connections

One common approach is to allocate a subnet that has not been assigned to hosts and to variably subnet it for use with connectivity between, rather than within, areas.

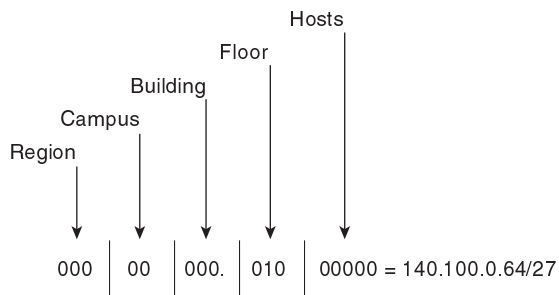
In reference to the case study described earlier, it would be sensible to take a subnet from the bits allocated to the buildings. Because there are enough bits allocated to address eight buildings, you have twice as many subnets as required. Even with the possibility of growth, one subnet would not be missed. Because the building bits come after the bits assigned to the campus, you must make a choice as to which campus will be selected for the honor of contributing a subnet of WAN addressing. This is an arbitrary decision that you need to document. If necessary, a building subnet can be commandeered from each campus.

If possible, the subnet you use should have nothing to do with any of the existing subnets. There is a consistency in numbering that identifies the WAN links, so in a troubleshooting environment, you can immediately see that a WAN link is causing the trouble and will not confuse the subnet (VLSM) with an existing segment.

In this example, if you use the bit pattern 000 as the network address for the building section, as well as for the campus and the region, the third octet would result in a 0. The network address for all interconnectivity would be 140.100.0. . . The last octet would be available for further subnetting with VLSM.

The subnet chosen for the WAN connections will be subnetted further using 30 bits of subnetting. This allows for only two hosts and is therefore a very efficient mask for point-to-point links.

Remember that the old rule for not using all 0s or all 1s is based on the entire subnet, not on the octet boundary. However, it is also important to remember that there is no longer a problem with subnet zero, which current Cisco IOS allows by default. Figure 2-8 shows assigning IP VLSM subnets for WAN connections.

Figure 2-8 *Assigning IP VLSM Subnets for WAN Connections*

The following is an example of how the addressing might be broken down.

Between the buildings in California:

- 140.100.0.64/27
- A 27-bit mask allows for 30 end-system addresses. This assumes that the buildings are connected via FDDI or Fast Ethernet.
- The range of hosts is 140.100.0.65 to 140.100.0.94.
- The broadcast address is 140.100.0.95.

Between the buildings and the campuses in California:

- 140.100.0.32/30
- 140.100.0.20/30
- 140.100.0.24/30
- 140.100.0.28/30

The prefix mask of /30 provides two host addresses, which allows for point-to-point addresses using Frame Relay.

Between the campuses and the regions:

- 140.100.0.48/30
- 140.100.0.4/30
- 140.100.0.12/30

The prefix mask of /30 provides two host addresses, which allows for point-to-point addresses that might also be using Frame Relay.

Between the regions:

- 140.100.0.96/30
- 140.100.0.16/30
- 140.100.0.8/30

The prefix mask of /30 provides two host addresses, which allows for point-to-point addresses that might also be using Frame Relay or dedicated serial leased lines.

NOTE In the instance of a subnet being used to address WAN connections, it might not be possible to summarize these networks. To summarize subnets, the subnets contained in the summary address must be contiguous; otherwise, the router is confused as to where to send the data. In a WAN environment, the connections might not be within a confined area, but scattered throughout the network.

The rules and conditions for creating a valid and appropriate IP addressing scheme for the network are complicated. Among other things, the addressing scheme must allow for growth, to scale over time. What works today might not be flexible for next year's business requirements. You cannot build a network that will accommodate every change and addition to its environment. With careful design, however, it might be possible to anticipate some of these changes and to ensure a network with enough flexibility to survive the changes.

Summarization

Having assigned IP addressing based on a hierarchical design, you can now consider the full weight of the advantages of VLSM in implementing summarization. The primary advantage is the reduction in network traffic and the size of the routing table.

Summarization allows the representation of a series of networks in a single summary address.

The reasons that the Internet implemented CIDR are equally pertinent in a single organization. VLSM and CIDR use the same principles, with VLSM being just an extension of CIDR at the organizational level.

At the top of the hierarchical design, the subnets in the routing table are more generalized. The subnet masks are shorter because they have aggregated the subnets lower in the network hierarchy. These summarized networks are often referred to as *supernets*, particularly when seen in the Internet aggregation of class addresses. They are also known as aggregated routes. Figure 2-9 shows the physical network design for the case study discussed earlier. Figure 2-10 shows the allocation of addresses using VLSM to support summarization for this network design.

Figure 2-9 *The Application of Summarized Routes on a Hierarchically Designed Network*

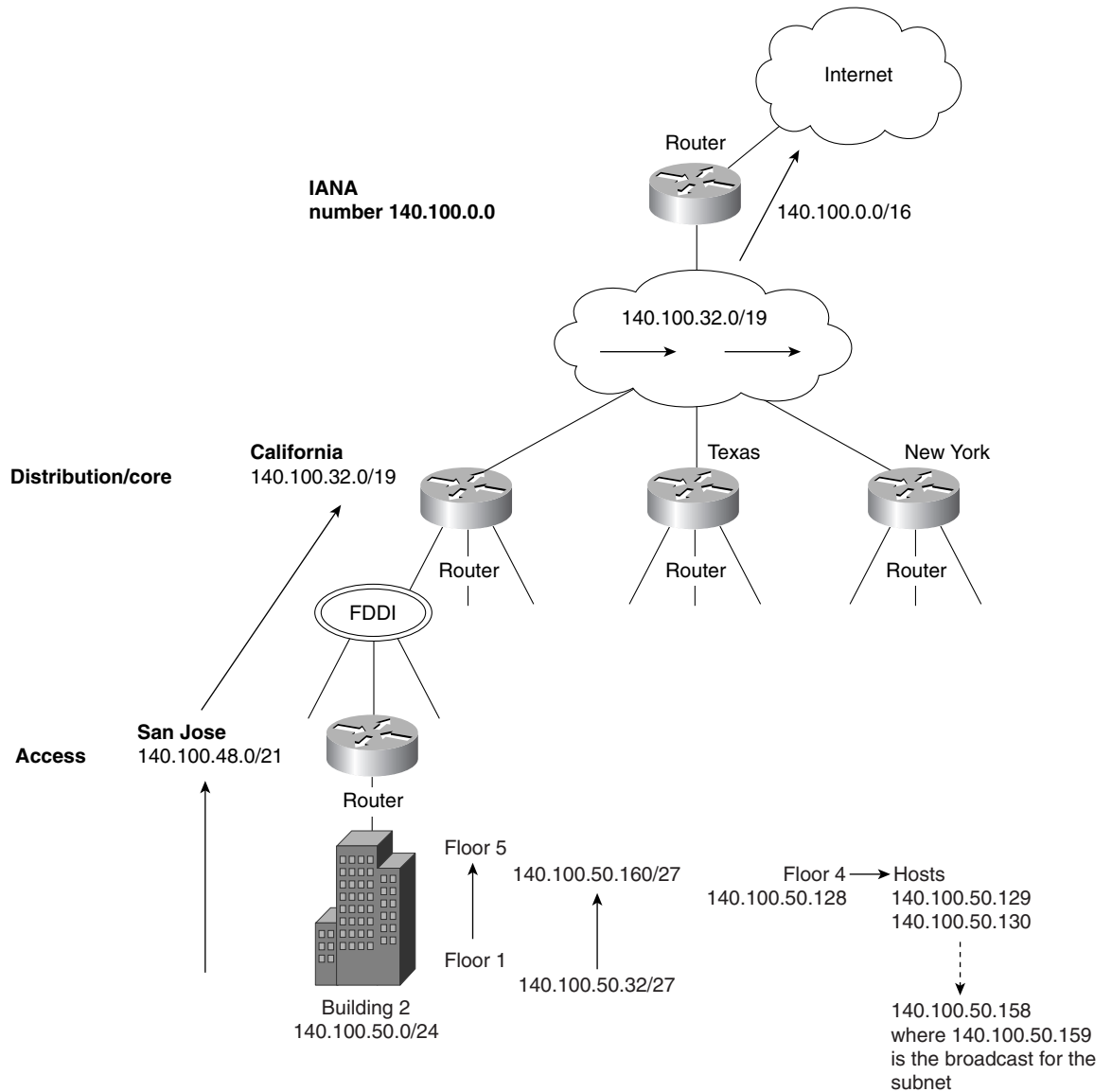
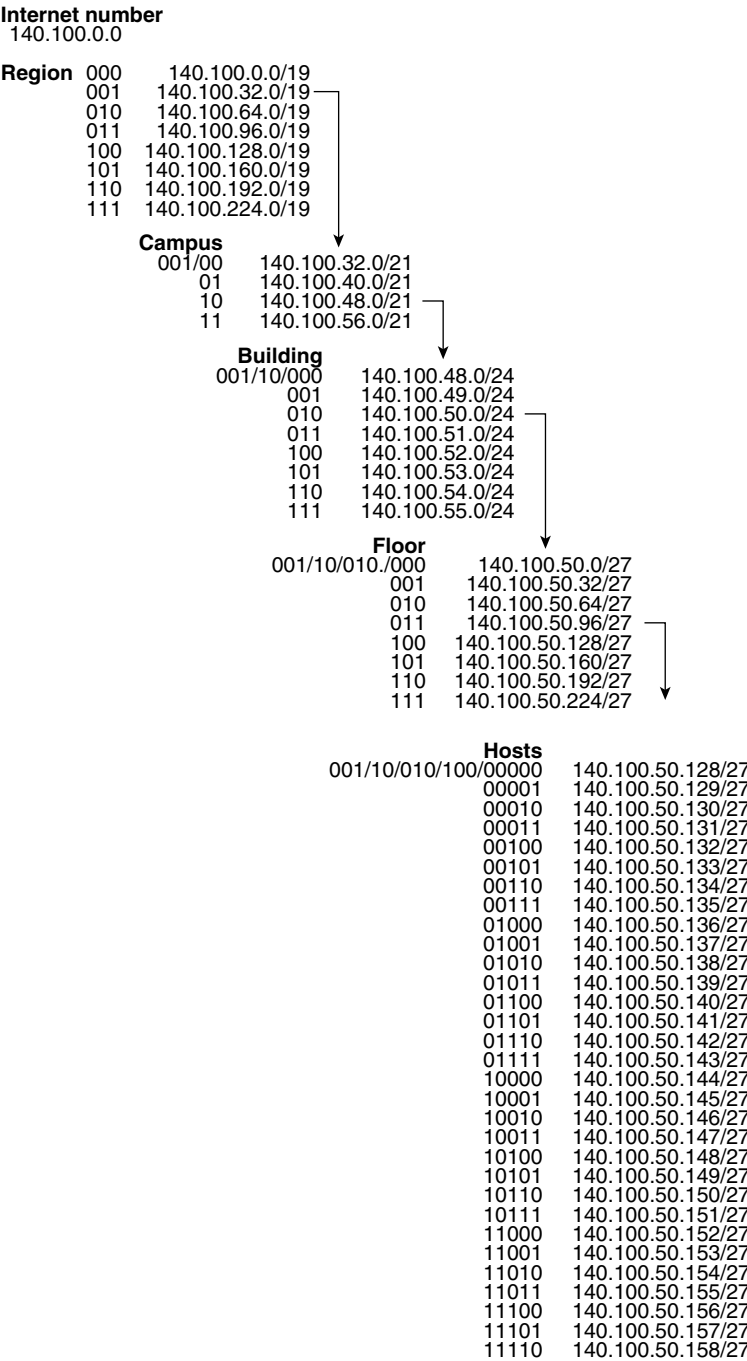


Figure 2-10 The Binary Calculation of the Hierarchical Addressing for the Organization



The Advantages of Summarization

The capability to summarize multiple subnets within a few subnets has the following advantages, as discussed in the next few sections:

- Reduces the size of the routing table
- Simplifies recalculation of the network
- Hides network changes
- Allows networks to grow

Reducing the Size of the Routing Table

In reducing the size of the routing table, the updates are smaller, demanding less bandwidth from the network. A smaller routing table also requires less memory in the router or CPU in the routing process itself because the lookup is quicker and more efficient.

The recalculation of the network is also simplified by maintaining small routing tables.

Hiding Network Changes

If the routing table contains a summary of the networks beneath it, any changes in the network at these levels are not seen. This is both a good thing and a bad thing. If the network in the earlier case study—140.100.50.128/27, the subnet on the fourth floor of the second building in San Jose, California—were to go down, the router at the core would be oblivious to the LAN problem. This is beneficial because there are no additional updates or recalculation.

The disadvantage is that any traffic destined for that subnet is sent on the assumption that it exists. To be more accurate, the core router sees the inbound IP packet destined for 140.100.50.131 and, instead of applying the /27 mask, uses the mask that it has configured. It employs the /19 mask that sees the subnet 140.100.32.0/19, although in reality the destination subnet is 140.100.50.128/27. If the subnet 140.100.50.128 is no longer available, all traffic is still forwarded until it reaches a router that sees the network 140.100.50.128 as directly connected or to the first router that sees the network 140.100.50.128 as unavailable. This would be a router using the /27 bit mask. An ICMP message that the network is unreachable is generated to the transmitting host. The host might stop transmitting after hearing that the network is down.

Although unnecessary traffic will traverse the network for a while, it is a minor inconvenience compared to the routing update demands on the network and the CPU utilization on the routers in large networks.

Other Solutions to Address Exhaustion

The efficient use of IP addressing, through prefix routing, CIDR, and VLSM, helps to alleviate address exhaustion experienced by the Internet; however, there are a few other methods that can be used. These are discussed in this section.

The use of the Cisco feature IP unnumbered is useful on the point-to-point serial lines because it saves the use of a subnet. *IP unnumbered* is a utility that allows point-to-point serial lines to have no IP address assigned. This is possible because the serial line is literally a pipe with two directly connected hosts. Each end of the serial line borrows an IP address from another interface on the Cisco router if an address is required: for example, when generating an IP packet and needing a source address for the packet header.

Cisco's use of secondary addressing is useful because it provides two subnets to a physical interface and, therefore, more available host bits. This does not save address space, but it is a solution for routing protocols that do not support VLSM. Some compatibility issues exist with some IP routing protocols; for example, not all routing protocols will see the second subnet.

Configuring Summarization

Summarization allows networks to grow because the network overhead can scale.

In the newer routing protocols, summarization must be manually configured; this manual configuration adds subtlety and strength. Each routing protocol deals with summarization in a slightly different way. How summarization works or is configured depends on the routing protocol used. This is discussed in Chapter 5, "IP Link-State Routing Principles."

NOTE Although Border Gateway Protocol (BGP) and Enhanced IGRP (EIGRP) perform automatic summarization, the summarization is done at the classful network boundary, using the first octet rule. This is the same as with older routing protocols, such as RIP.

Automatic Summarization

All routing protocols employ some level of summarization. The older protocols, such as RIP and IGRP, automatically summarize at the Internet address or natural class boundary. They have no choice because the subnet mask is not sent in the routing updates. When a routing update is received, the router looks to see whether it has an interface in the same classful network. If it has one, it applies the mask configured on the interface to the incoming routing update. With no interface configured in the same Internet address, there is insufficient information and the routing protocol uses the natural mask for the routing update. Automatic summarization uses the first octet rule.

Manual Summarization

EIGRP, IS-IS, RIPv2, and OSPF are more sophisticated. They send the subnet mask along with the routing update. This feature allows the use of VLSM and manual summarization. When the routing update is received, it assigns the mask to the particular subnet. When the routing process performs a lookup, it searches the entire database and acts on the longest match. Searching the routing table for the longest match is an important feature because it allows the following:

- The granularity of the hierarchical design
- Manual summarization
- Discontiguous networks

Discontiguous Networks

A *discontiguous network* refers to a network in which a different classful network separates two instances of the same classful network. This can happen through either intentional design or a break in the network topology. If the network is not using a routing protocol that supports VLSM, this creates a problem, because the router does not know where to send the traffic. Without a subnet mask, it resolves the address down to the classful network, which appears as if there is a duplicate address. The same classful network appears twice, but in different locations. In most cases, the router will load balance between the two paths leading to the two instances of the one classful network address, the two discontiguous subnets. As with any multiple entry in a routing table, the router will load balance over the multiple paths if they are equal, resulting in only a portion of the traffic taking the correct path. The symptoms that the network will see are those of intermittent connectivity.

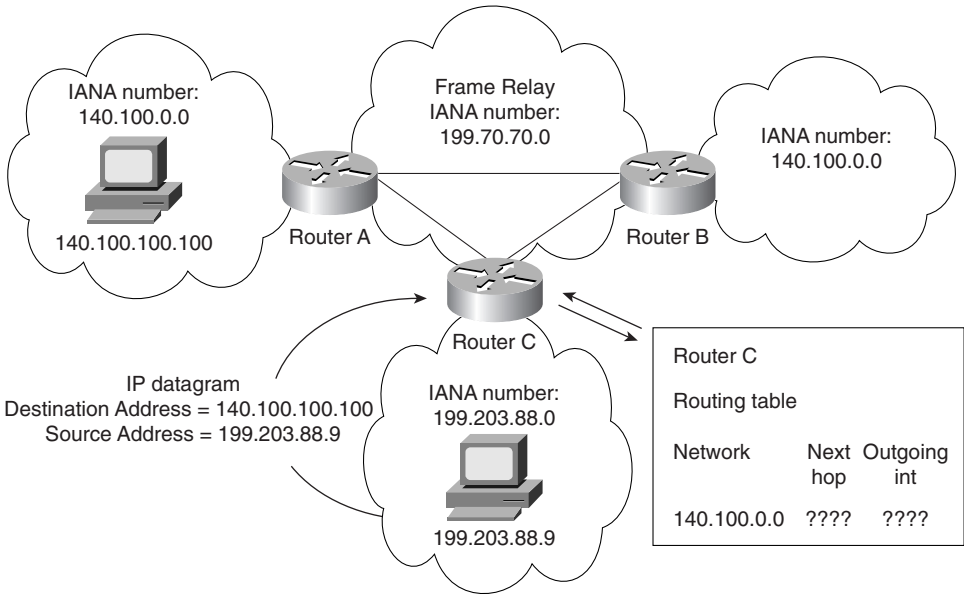
Figure 2-11 shows an instance of a discontiguous network.

Considerations for Summarization with Discontiguous Networks

Discontiguous networks are not a problem with VLSM, because the routing table does a lookup based on the longest match; therefore, the routing process will choose the network with the longest mask and no duplicate path is seen. However, if VLSM is used on networks that employ automatic summarization, problems of discontiguous networks could arise. Despite the fact that VLSM can distinguish between network 131.108.16.0/20 and 131.108.20.0/24, automatic summarization would reduce these separate networks to 131.108.0.0. If these networks are separated by another classful network, it would cause discontiguous network problems.

Manual summarization allows the administrator to create summarization with greater granularity and thus avoid such problems. Also, if a hierarchical design has been implemented, it is possible that discontiguous networks will not arise when summarization is used, as 131.108.20.0 would be a smaller branch off the main branch of 131.108.16.0.

Figure 2-11 *Discontiguous Networks*



If there are discontiguous networks in the organization, it is important that summarization is turned off or not configured. Summarization might not provide enough information to the routing table on the other side of the intervening classful network to be capable of appropriately routing to the destination subnets. This is especially true of EIGRP, which automatically summarizes at the classful network boundary, which would be disastrous in this situation.

In OSPF and EIGRP, manual configuration is required for any sophistication in the network design. It is not always possible to achieve summarization because it depends entirely on the addressing scheme that has been deployed. However, because EIGRP can perform summarization at the interface level, it is possible to select interfaces that do not feed discontiguous networks for summarization. This capability to summarize selectively is very powerful.

The key to whether summarization is configurable is determined by whether there are common high-order bits in the addresses.

As demonstrated in the case study “Addressing the Network” earlier in this chapter, the design has created common high-order bits to facilitate summarization. The addressing scheme for the case study, shown in Figure 2-7, shows that every campus within a region will share the same high-order bits (those to the left). In California, every campus, building, floor, and host will share the bits 001, whereas within the California campus of San Jose, every building shares the high-order bits of 001 10. Therefore, it is very simple to configure summarization.

This is not necessarily the case if the addressing structure is already in place. Some analysis of the addressing scheme is required to decide whether summarization can be configured.

If summarization is deemed impossible, you have the following two options:

- Don't summarize, but understand the scaling limitations that have now been set on the network.
- Readdress the network. This task is not to be underestimated, although the advantages may well make it worthwhile.

Foundation Summary

The “Foundation Summary” section of each chapter lists the most important facts from the chapter. Although this section does not list every fact from the chapter that will be on your exam, a well-prepared candidate should, at a minimum, know all the details in each “Foundation Summary” before going to take the exam.

Remember the following important points regarding IP addressing:

- The IP address is 32 bits long.
- The network/host boundary can be anywhere in the 32 bits.
- The Internet allocates a unique bit pattern. These bits are the first bits on the far left and are not available for you to use for networks because they identify your organization to the Internet.
- The Internet authority ARIN (in the United States) will provide the left (network) portion of the address to the organization to use for addressing within their network. The bits provided on the right portion of the address are zeroed and are allocated by the organization as subnet and host addresses.
- The network mask is the identification of the bits allocated to the network, defined on all participating routers.

CIDR solves the problem of the excessive network resources required to manage the huge routing tables. Table 2-7 and Table 2-8 include further quick reference CIDR material.

Table 2-7 *RFCs about CIDR*

RFC Number	Title
1517	Applicability Statement for the Implementation of Classless InterDomain Routing (CIDR)
1518	An Architecture for IP Address Allocation with CIDR
1519	Classless Interdomain Routing (CIDR): An Address Assignment and Aggregation Strategy
1520	Exchanging Routing Information Across Provider Boundaries in the CIDR Environment

Table 2-8 *Table to Illustrate the Use of Prefix Masks*

Prefix	Mask	New Address Space
/27	255.255.255.224	12 percent of Class C 30 hosts
/26	255.255.255.192	24 percent of Class C 62 hosts
/25	255.255.255.128	50 percent of Class C 126 hosts
/23	255.255.254.0	2 Class Cs 510 hosts
/22	255.255.252.0	4 Class Cs 1022 hosts
/21	255.255.248.0	8 Class Cs 2046 hosts
/20	255.255.240.0	16 Class Cs 4094 hosts

Table 2-9 shows an example of an IP address in both decimal and binary format.

Table 2-9 *An IP Address and Mask Shown in Binary*

Description	Octet 1	Octet 2	Octet 3	Octet 4
IANA address in decimal	200	100	48	0
IANA address in binary	11001000	01100100	00110000	00000000
Prefix as a subnet mask in decimal	255	255	248	0
Prefix as a subnet mask in binary	11111111	11111111	11111000	00000000

If it were a standard Class C address, the mask would be 255.255.255.0. By making the mask 255.255.248.0, the last three bits of the third octet are essentially giving the organization eight Class C networks.

The following main rules apply when subnetting:

- A subnet can be used to address hosts, or it can be used for further subnetting.
- All 1s or all 0s in the subnet portion of the classful network cannot be used. If this rule has been followed, any subnet that is further subnetted does not need to obey this rule, as it has already been observed.
- The routing protocol must carry the subnet mask in its updates.
- Multiple IP addresses intended to be summarized must have the same high-order bits.
- Routing decisions are made on the entire address, preferring the longest bit pattern available.

The two main reasons for using VLSM are as follows:

- To make efficient use of the available addressing
- To enforce a good hierarchical design, allowing summarization and documentation

The advantages of summarization are as follows:

- To reduce the size of the routing table.
- To reduce network overhead.
- To make communication of routing updates more efficient.
- To reduce CPU and memory utilization.
- To simplify management.
- To maximize the use of IP addresses by allowing a more granular application of addresses. Thus, a point-to-point link can be given two host addresses, while a switched LAN has 254 host addresses available. Otherwise, the point-to-point link would be assigned the equivalent of a Class C address.
- To isolate topographical changes from other areas.

Q&A

As mentioned in the introduction, “All About the CCNP, CCDP, and CCIP Certifications,” you have two choices for review questions. The questions that follow next give you a bigger challenge than the exam itself by using an open-ended question format. By reviewing now with this more difficult question format, you can exercise your memory better and prove your conceptual and factual knowledge of this chapter. The answers to these questions are found in Appendix A.

For more practice with examlike question formats, including questions using a router simulator and multichoice questions, use the exam engine on the CD.

1. Identify one criterion to help determine a subnet mask for classless addressing when designing a network-addressing scheme.
2. With a classless address of 204.1.64.0/20, what is the range of classful addresses that are included in the address? Write your answer in dotted decimal and the third octet in binary notation.
3. What is a discontinuous network?
4. For VLSM to be available as a design option in the network, what characteristic must the routing protocol possess?
5. If summarization is to be implemented in the network, name one design criterion for the addressing scheme that must be in place.
6. If the host portion of a subnet has been used to identify end devices, can that subnet be used again for VLSM?
7. Give one example of when route summarization would not be a good solution.
8. Give one reason for implementing route summarization.
9. Given an address of 133.44.0.0 and a prefix mask of /25, how many networks can be addressed, and how many hosts can exist on each network? Write the first and last possible subnets in binary and decimal notation.
10. What class of address is 131.188.0.0, and how many hosts can be addressed if no subnetting is used?
11. Write out the decimal notation of the following subnet mask presented in the binary notation of 11111111.11111111.11111111.11111000.
12. Is 201.111.16.0/20 a valid subnet mask?
13. Briefly define route summarization.

- 14.** What sort of design scheme does route summarization require?
- 15.** In route summarization, in which direction is the network/host boundary in the subnet mask moved?
- 16.** Explain how summarization allows for smaller routing tables?
- 17.** What is the subnet mask for a /21 prefix?
- 18.** What is the default subnet mask for the IP address 192.18.16.15?
- 19.** State whether 131.104.0.0/13 is an example of CIDR or VLSM routing.
- 20.** State how many classful addresses are summarized in the address 131.104.0.0/13.

Scenarios

The following scenarios and questions are designed to draw together the content of the chapter and exercise your understanding of the concepts. There is not necessarily a right answer. The thought process and practice in manipulating the concepts is the goal of this section. The answers to the scenario questions are found at the end of this chapter.

Scenario 2-1

A network has a remarkably even distribution of campuses, buildings, and hosts. The company has four campuses, each campus has four buildings, each building has five floors, and each floor has approximately 100 hosts. Each building also has a basement where the building servers are held.

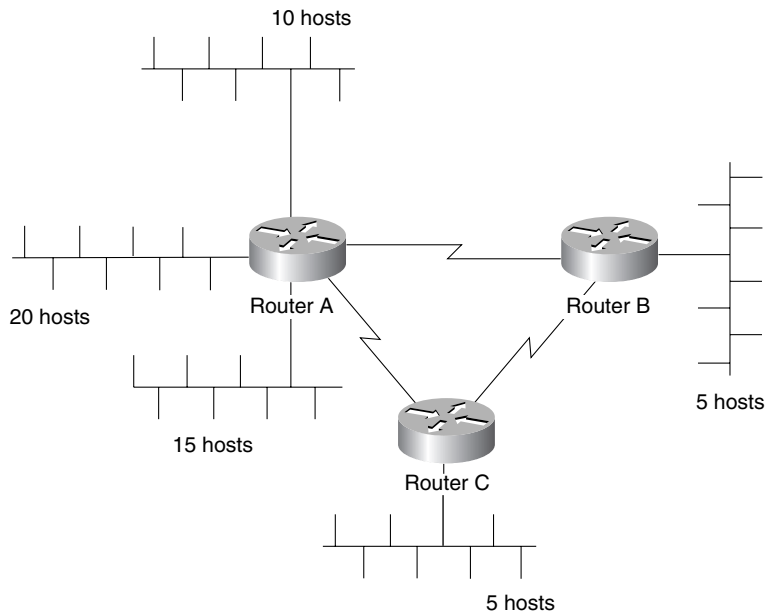
There are eight locations distributed globally. Each location replicates this physical design. The locations are connected via dedicated leased T1 lines. Each T1 constitutes a subnet.

1. Draw the topology map for one of the locations.
2. Using the network address 10.0.0.0, design an addressing scheme that can be summarized. Apply the binary notation for the bit allocation to your diagram.
3. List the range of hosts on one of the subnets allocated to a floor in a building.
4. Indicate how summarization would work within the location.
5. Allocate a subnet to be used for VLSM to address the WAN links between the locations.
6. Is it possible to summarize the WAN subnets?

Scenario 2-2

Study Figure 2-12, and answer the questions that follow.

Figure 2-12 *Topology Map for Scenario 2-2*



1. To address this network, what class of address would you apply to the Internet?
2. Could you use VLSM? Give reasons for your answer.
3. If you could use VLSM, write out the masks that you would deploy in binary notation.
4. Could summarization be implemented?

Scenario Answers

The answers provided in this section are not necessarily the only possible answers to the questions. The questions are designed to test your knowledge and to give practical exercise in certain key areas. This section is intended to test and exercise skills and concepts detailed in the body of this chapter.

If your answer is different, ask yourself whether it follows the tenets explained in the answers provided. Your answer is correct not if it matches the solution provided in the book, but rather if it has included the principles of design laid out in the chapter.

In this way, the testing provided in these scenarios is deeper: It examines not only your knowledge, but also your understanding and capability to apply that knowledge to problems.

If you do not get the correct answer, refer back to the text and review the subject tested. Be certain to also review your notes on the question to ensure that you understand the principles of the subject.

Scenario 2-1 Answers

Using the network address 10.0.0.0, designing a summarized addressing scheme is straightforward.

When the last three octets are written in binary notation, it is easy to determine the bit allocation needed to fulfill the requirements.

Location	Campus	Building	Floor	Hosts
0000	0000	.0000	0000	.00000000

This design provides 16 locations, 16 campuses, and 16 buildings. This would allow 254 hosts per floor or building subnet. Therefore, there is a lot of flexibility in this design for future growth.

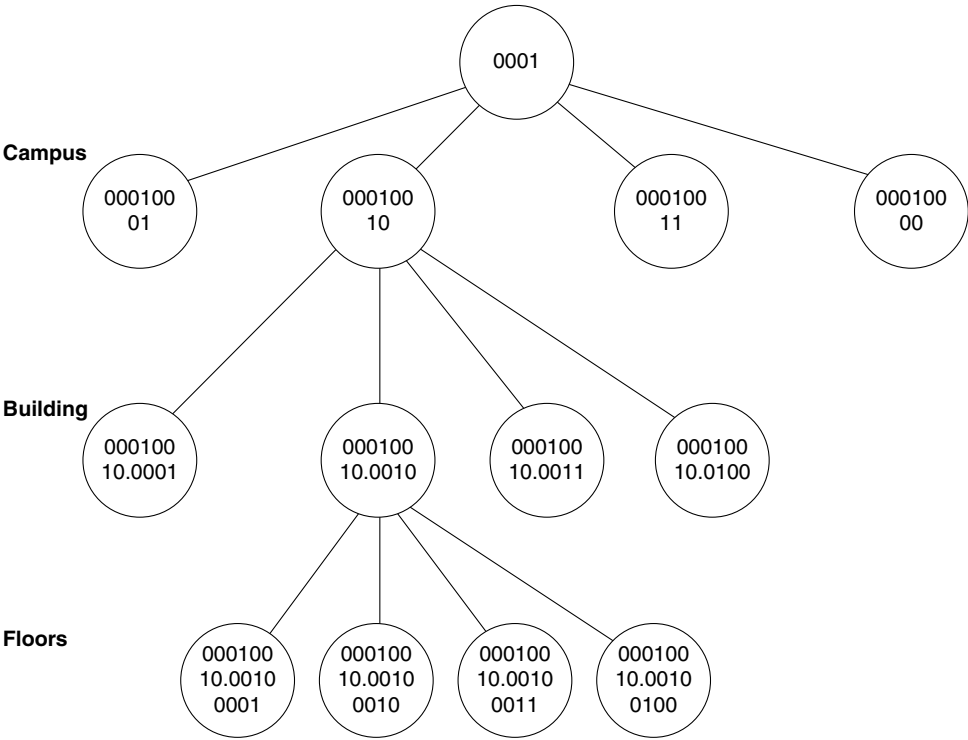
1. *Draw the topology map for one of the locations.*

See Figure 2-13.

2. *Using the network address 10.0.0.0, design an addressing scheme that can be summarized. Apply the binary notation for the bit allocation to your diagram.*

See Figure 2-13.

Figure 2-13
 Topology Map of One of the Locations



3. List the range of hosts on one of the subnets allocated to a floor in a building.

If one of the floors is given the subnet:

Location	Campus	Building	Floor	Hosts
0000	0000	.0000	0000.	.00000000
Subnet in binary notation:				
00001010.00010010.00100010.00000000				
Subnet in decimal notation:				
10.18.34.0				
Range of hosts on that subnet:				
10.18.34.1 to 10.18.34.254				

4. *Indicate how summarization would work within the location.*

Summarization would work within a location because every device and subnet would share the same four high-order bits of the second octet. The following example demonstrates this:

Location	Campus	Building	Floor	Hosts
Using network 10.0.0.0, the four high-order bits of the second octet identify the location.				
00001010.0000	0000	.0000	0000	.00000000
Subnet in binary notation: 00001010.0001 0010 .0010 0010 .00000000				
Subnet in decimal notation: 10.18.34.0				

The summarized address advertised out of the location router would be 10.16.0.0/12.

5. *Allocate a subnet to be used for VLSM to address the WAN links between the locations.*

Many spare subnets are available in the addressing scheme designed. To address the WAN links, it would be sensible to select one of the subnets allocated to the floors and to reassign it to be further subnetted. For example:

Location	Campus	Building	Floor	Hosts
0000	0000	.0000	0000.	.00000000
Subnet in binary notation: 10000000.10001000.00000000				
Subnet in decimal notation: 10.128.128.0/30				

This allows 64 subnets, with each subnet allowing two hosts (ideal for point-to-point lines). The use of 128.128 in the second and third octets eases network management by readily identifying the serial connections.

6. *Is it possible to summarize the WAN subnets?*

It would not be easy to summarize these WAN subnets because they have a longer bit pattern than the other subnets beneath them. If summarization is possible, they could be summarized down to 10.128.128.0. It is equally sensible to use any easily recognizable address for WAN links (for example, 10.100.100.0).

Scenario 2-2 Answers

1. *To address this network, what class of address would you apply to the Internet?*

You can address the network using a Class C address.

2. *Could you use VLSM? Give reasons for your answer.*

You can use VLSM as long as you use a routing protocol to support the propagation of the subnet mask. It would be useful to have VLSM for the WAN links, but not essential.

3. *If you could use VLSM, write out the masks that you would deploy in binary notation.*

The bit allocation of the fourth octet could be as follows:

Remote Subnet Locations	Hosts
000	00000

This would allow for six remote subnet locations, with 30 hosts on each subnet. The assumption was that the company was more likely to expand each existing location than to increase the number of remote sites. If the reverse were true, the mask would no longer be appropriate, and a single Class C may no longer be sufficient.

Because there are only three remote sites, with five networks to address and three WAN point-to-point links, and because there are six available subnets, one of the subnets could be further subnetted. This subnet would be used to address the WAN links. Another alternative is to use ip unnumbered on the serial links.

Remote Subnet Locations	Hosts
000	00000
110	00000 taken for WAN links
New mask:	
110000	00

This would allow 14 WAN links to be identified.

NOTE It is possible to address more links with the use of subnet zero.

4. *Could summarization be implemented?*

In this size of a network, summarization is not a concern and would not be possible; also, there is no hierarchy in the physical design.

It should be noted that this design does not allow for much network growth, and the organization might want to consider using a private Class B network. Private addressing is discussed in Chapter 3.



This chapter covers the following topics, which you need to understand to pass the CCNP/CCDP/CCIP BSCI exam:

- Criteria in designing IP networks
- Private addresses on the Internet
- Connecting to the outside world with NAT
- Understanding IPv6

Designing IP Networks

This chapter deals with designing IP networks to efficiently use the addressing structure to reduce the routing tables and thus conserve network resources.

Many of the design principles for an IP network were dealt with in Chapter 2, “IP Addressing.” The approach in Chapter 2 was a practical one, explaining how to address a network and the need for a hierarchical design. By contrast, this chapter examines the design criteria from a high-level perspective. Instead of describing how you would design a hierarchical addressing scheme to support variable-length subnet masks (VLSM) and allow summarization, this chapter explains why there is a need for hierarchical structure and summarization.

This chapter also considers the relatively recent solution to the problem of applying for limited Internet addresses from the IANA, private addressing. When using private addresses, any connection to the Internet must be handled in such a way that duplicate addresses do not appear in the public domain. Duplicate addresses result in a lack of connectivity. Solutions to the problem of duplicate addressing are explored within this chapter.

IPv6 and its features are discussed in terms of network design and implementation. In particular, the advantages of IPv6 over IPv4 are considered. The IPv6 addressing format is also explained, as are the various methods of transitioning an IPv4 network to run IPv6 and the routing protocols that are available for the new IP stack.

“Do I Know This Already?” Quiz

The purpose of the “Do I Know This Already?” quiz is to help you decide what parts of this chapter to use. If you already intend to read the entire chapter, you do not necessarily need to answer these questions now.

The 15-question quiz, derived from the major sections in the “Foundation Topics” portion of the chapter, helps you to determine how to spend your limited study time.

Table 3-1 outlines the major topics discussed in this chapter and the “Do I Know This Already?” quiz questions that correspond to those topics.

Table 3-1 “Do I Know This Already?” Foundation Topics Section-to-Question Mapping

Foundation Topics Section	Questions Covered in This Section
Criteria in Designing IP Networks	1–6
Private Addresses on the Internet	7–9
Connecting to the Outside World with NAT	9–12
Understanding IPv6	13–15

NOTE The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark this question wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

- Which of the following are valid uses for access lists?
 - Restricting networks sent out in routing updates
 - Restricting connectivity to remote networks
 - Preventing users interactive access to the routing table
 - Restricting large packet sizes from traversing the network
- In the hierarchical design suggested by Cisco, at which layer are access lists not recommended?
 - Core layer
 - Access layer
 - Distribution layer
 - Access lists are recommended at all layers
- What is the function of the distribution layer?
 - Connects the routers to the end systems, allowing the datagrams to be distributed to the destination address
 - Distributes the datagrams to the rest of the network. It is the pinnacle of the network, connecting the other layers to one another.
 - Provides the demarcation point between the core and access layers, providing policy-based connectivity and allowing you to do packet manipulation

- d. Provides the connections to other autonomous systems or to the Internet, distributing data to the outside world
4. What feature is required in a routing protocol to allow it to support prefix routing?
 - a. Interoperability with other vendors
 - b. The extended prefix or subnet mask is sent in the routing update.
 - c. Hierarchical addressing
 - d. The protocol must be VLSM compliant.
 5. In order for a network to support summarization, which of the following must be true?
 - a. VLSM needs to be configured on the DHCP server.
 - b. The addresses to be summarized share the same high order bits.
 - c. The physical and logical topology of the network is hierarchical.
 - d. The design of the network must conform to the rules laid out in RFC 2011.
 6. In designing the IP addressing of a network, which of the following questions are pertinent to the design process?
 - a. How many subnets exist currently in your network?
 - b. Is access to the subnets required from other subnets?
 - c. Are you running PSCA?
 - d. Where are the subnets in relation to the topology map?
 7. What are the private addresses allocated in RFC 1918?
 - a. Class A: 10.0.0.0
 - b. Class A: 10.0.0.0, Class B: 172.16.0.0, Class C: 192.168.1.0
 - c. Class A: 10.0.0.0, Class B: 172.16.0.0–172.32.0.0, Class C: 192.168.1.0–192.168.254.0
 - d. Class A: 10.0.0.0, Class B: 172.16.0.0–172.32.0.0, Class C: 192.168.0.0–192.168.255.0
 8. What is the purpose of private addressing?
 - a. To allow companies to have no communication with the Internet
 - b. To allow companies to address the networks within their autonomous system without any constraint on limited IP address space on the Internet
 - c. To configure encryption on individual end systems
 - d. To prevent spam attacks

9. Is it necessary to prevent the private addresses from entering the Internet?
 - a. Yes, if the private addresses appear on the Internet, the packet with the address should be dropped.
 - b. Yes, the private address would invite unsolicited mail.
 - c. No, the private address is unique to the Internet; it simply indicates that there is a firewall in place.
 - d. No, the destination address will convert the source address to a unique address.
10. Which NAT feature would you implement if you wished to create a one-to-one translation?
 - a. Dynamic Source Address Translation
 - b. Port Address Translation
 - c. Static Addressing
 - d. Destination address rotary translation
11. Which of the following most closely defines a feature of NAT?
 - a. To translate an address on one network into a different address for another network
 - b. To encrypt the source address to prevent spam attacks
 - c. To translate private addresses for use in a WINS server
 - d. To store prefix routing information on the Internet
12. Which of the following is a feature of NAT that is supported by Cisco?
 - a. Encryption
 - b. Destination Allocation Resource Translation
 - c. Port Allocation Translation
 - d. Port Address Translation
13. Which of the following is a valid IPv6 address?
 - a. 4021::240E::0AC0:3428:121C
 - b. 4021:240E::0AC0:3428::
 - c. 4021::0000::240E::0000::0000::0AC0::3428::121C
 - d. 4021:0:240E::0AC0:3428:121C

14. How long is an IPv6 address?
 - a. 16 hexadecimal numbers
 - b. 32 decimal numbers
 - c. 128 bits
 - d. 32 bits

15. The multicast address range for IPv6 is which of the following?
 - a. F000::/8 – FF0F::/8
 - b. FF00::/8 – FFFF::/8
 - c. F000::/8 – FFFF::/8
 - d. FF00::/8 – FF0F::/8

The answers to the “Do I Know This Already?” quiz are found in Appendix A, “Chapter ‘Do I Know This Already?’ Quizzes and Q&A Sections.” The suggested choices for your next step are as follows:

- **8 or less overall score** —Read the entire chapter. This includes the “Foundation Topics” and “Foundation Summary” sections, the “Q&A” section, and the “Scenarios” at the end of the chapter.
- **9–12 overall score** —Begin with the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. If you have trouble with these exercises, read the appropriate sections in “Foundation Topics.”
- **13 or more overall score** —If you want more review on these topics, skip to the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. Otherwise, move to the next chapter.

Foundation Topics

Criteria in Designing IP Networks

The topics in this section are important both for understanding IP addressing and routing and for use within the context of the Cisco certification. The need for the hierarchical design proposed by Cisco is discussed, explaining the function of each layer and how access lists are used in conjunction with this design to protect the network from excessive and redundant traffic.

This section deals with the hierarchical design that Cisco uses. The design integrates well with VLSM design because summarization helps to ensure a stable and well-managed network. This section also includes a review of access lists and their use, because they are integral to IP network design. You will examine alternatives to access lists and identify other key points to remember when designing an IP network.

The Cisco Hierarchical Design

It is crucial to build a network that can grow or scale with the needs of the user. This avoids a network that reels from crisis to crisis. Cisco provides a hierarchical design that simplifies network management and also allows the network to grow. This growth may be physical growth or capacity growth.

To achieve a stable and responsive network—and to keep local traffic local, preventing network congestion—Cisco suggests a network design structure that allows for growth. The key to the design is making it hierarchical, with a division of functionality between the layers of the hierarchy. Traffic that begins on a lower layer of the hierarchy is only allowed to be forwarded through to the upper levels if it meets clearly defined criteria. A filtering operation restricts unnecessary traffic from traversing the entire network. Thus, the network is more adaptable, scalable, and reliable.

Clear guidelines and rules govern how to design networks according to these principles. The following section explains how the hierarchical network design proposed by Cisco reduces congestion.

If the network is designed hierarchically, with each layer acting as a filter for the layer beneath it, the network can grow effectively. In this way, local traffic is kept local (within the same layer), and only data and information about global resources needs to travel outside the immediate domain or layer.

Understanding that the layers are filtering functions begs the question of how many layers are required in your network. The answer is that it depends on the type of applications and network architecture, in addition to other criteria.

The Cisco design methodology is based on simplicity and filtering. Cisco suggests that the largest networks currently require no more than three layers of filtering.

Because a hierarchical layer in the network topology is a control point for traffic flow, a hierarchical layer is the same as a routing layer. Thus, a layer of hierarchy is created with the placement of a router or a Layer 3 switching device.

The number of hierarchical layers that you need to implement in your network reflects the amount of traffic control required. To determine how many layers are required, you must identify the function that each layer will have within your network.

The Functions of Each Layer

Each hierarchical layer in the network design is responsible for preventing unnecessary traffic from being forwarded to the higher layers, only to be discarded by unrelated or uninterested hosts. The goal is to allow only relevant traffic to traverse the network and thereby reduce the load on the network. If this goal is met, the network can scale more effectively. The three layers of a hierarchy are as follows:

- The access layer
- The distribution layer
- The core layer

The next sections describe each layer in more detail.

The Access Layer

In accordance with its name, the access layer is where the end devices connect to the network—where they gain access to the company network. The Layer 3 devices (such as routers) that guard the entry and exit to this layer are responsible for ensuring that all *local* server traffic does not leak out to the wider network. Quality of service (QoS) classification is performed here, along with other technologies that define the traffic that is to traverse the network. Service Advertisement Protocol (SAP) filters for NetWare and AppleTalk's GetZoneLists are also implemented here, in reference to the design consideration of client/server connectivity.

The Distribution Layer

The distribution layer provides connectivity between several parts of the access layer. The distribution layer is responsible for determining access across the campus backbone by filtering out

unnecessary resource updates and by selectively granting specific access to users and departments. Access lists are used not just as traffic filters, but as the first level of rudimentary security.

Access to the Internet is implemented here, requiring a more sophisticated security or firewall system.

The Core Layer

The responsibility of the core layer is to connect the entire enterprise by interconnecting distribution layer devices. At the pinnacle of the network, reliability is of the utmost importance. A break in the network at this level would result in the inability for large sections of the organization to communicate. To ensure continuous connectivity, the core layer should be designed to be highly redundant, and as much as possible, all latency should be removed. Because latency is created when decisions are required, decisions relating to complex routing issues, such as filters, should not be implemented at this layer. They should be implemented at the access or distribution layers, leaving the core layer with the simple duty of relaying the data as fast as possible to all areas of the network. In some implementations, QoS is implemented at this layer to ensure a higher priority to certain packets, preventing them from being lost during high congestion periods.

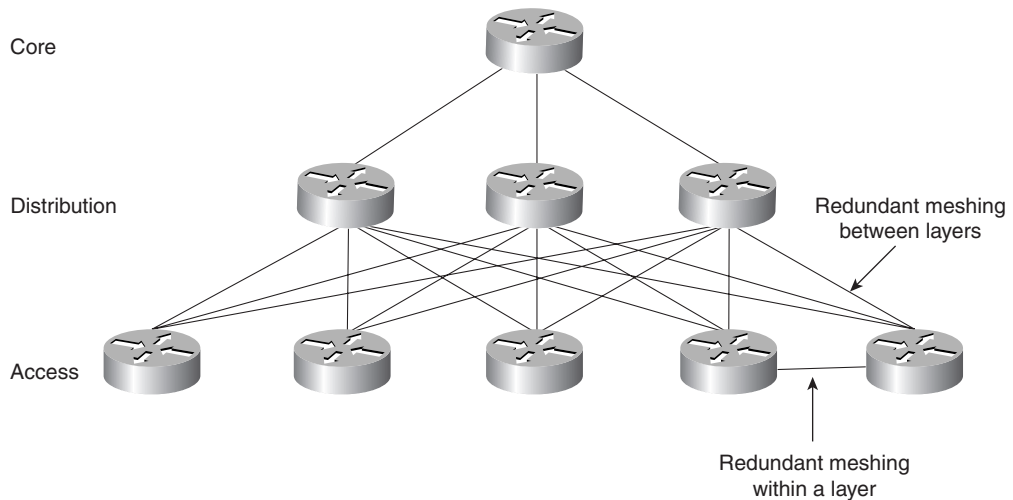
General Design Rules for Each Layer

A clear understanding of the traffic patterns within the organization—who is connecting to whom and when—helps to ensure the appropriate placement of client and servers, and eases the implementation of filtering at each layer. Without hierarchy, networks have less capacity to scale because the traffic must traverse every path to find its destination, and manageability becomes an issue.

It is important for each layer to communicate only with the layer above or below it. Any connectivity or meshing within a layer impedes the hierarchical design.

Organizations often design their networks with duplicate paths. This is to build network resilience so that the routing algorithm can immediately use an alternative path if the primary link fails. If this is the design strategy of your company, care should be taken to ensure that the hierarchical topology is still honored.

Figure 3-1 shows an illustration of the appropriate design and traffic flow.

Figure 3-1 *Redundant Connections Between Layers*

You need to have an understanding of the current network, the placement of the servers, and traffic flow patterns before attempting to design an improved network with the proper hierarchy.

One of the strengths of the Cisco hierarchical design is that it allows you to identify easily where to place the access lists. A quick review of access lists and how they can be used is provided in the next section.

IP Access Lists

Cisco router features enable you to control traffic, primarily through access lists. They are crucial to the sophisticated programming of a Cisco router and allow for great subtlety in the control of traffic.

Given that the router operates at Layer 3, the control that is offered is extensive. The router can also act at higher layers of the OSI model. This proves useful when identifying particular traffic and protocol types for prioritization across slower WAN links.

You can use access lists to either restrict or police traffic entering or leaving a specified interface. They are also used to implement “what if” logic on a Cisco router. This gives you the only real mechanism of programming the Cisco router. The access lists used for IP in this way enable you to apply subtlety to the router’s configuration. This section reviews how to configure access lists and discusses their use in an IP network. The books *CCNA Self-Study: Interconnecting Cisco Network Devices (ICND)* and the *CCNA ICND Exam Certification Guide (CCNA Self-Study, exam #640-811)*, both from Cisco Press, deal with these subjects in more depth.

Because access lists can be used so subtly in system programming, they are used in many ways. IP access lists are used mainly to manage traffic. The next sections discuss the role of access lists in security and controlling terminal access.

Security Using Access Lists

Cisco recommends using alternative methods rather than access lists for security. Although access lists are complex to conceive and write, they are easy to spoof and break through. As of IOS software version 11.3, Cisco implemented full security features. Use these features instead of access lists. The Cisco Secure Integrated Software (IOS Firewall Feature Set) is also now available.

Some simple security tasks are well suited to access lists, however. Although access lists do not constitute complex security, they will deter the idle user from exploring the company network.

The best way to use access lists for security is as the first hurdle in the system, to alleviate processing on the main firewall. Whether the processing on the firewall device is better designed for dealing with the whole security burden, or whether this task should be balanced between devices, should be the topic of a capacity-planning project.

Controlling Terminal Access

Access lists applied to router interfaces filter traffic traversing the router; they are not normally used to filter traffic generated by the router itself. To control Telnet traffic in which the router is the end station, an access list can be placed on the vty.

Five terminal sessions are available: vty 0 through vty 4. Because anticipating which session will be assigned to which terminal is difficult, control is generally placed uniformly on all virtual terminals. Although this is the default configuration, some platforms have different limitations on the number of vty interfaces that can be created.

Traffic Control Through Routing Updates

Traffic on the network must be managed. Traffic management is most easily accomplished at Layer 3 of the OSI model. You must be careful, however, because limiting traffic also limits connectivity. Therefore, careful design and documentation is required.

Routing updates convey information about the available networks. In most routing protocols, these updates are sent out periodically to ensure that every router's perception of the network is accurate and current.

Distribute Lists

Access lists that are applied to routing protocols restrict the information sent out in the update and are called *distribute lists*. Distribute lists work by omitting the routing information about certain networks based on the criteria in the access list. The result is that remote routers that are unaware of these networks are not capable of delivering traffic to them. Networks hidden in this way are typically research-and-development sites, test labs, secure areas, or just private networks. This is also a way to reduce overhead traffic in the network.

These distribute lists are also used to prevent routing loops in networks that have redistribution between multiple routing protocols.

When connecting two separate routing domains, the connection point of the domains, or the entry point to the Internet, is an area through which only limited information needs to be sent. Otherwise, routing tables become unmanageably large and consume large amounts of bandwidth.

Other Solutions to Traffic Control

Many administrators tune the update timers between routers, trading currency of information for optimization of bandwidth. All routers running the same routing protocol expect to hear these updates with the same frequency that they send out their own. If any of the parameters defining how the routing protocol works are changed, these alterations should be applied consistently throughout the network; otherwise, routers will time out and the routing tables will become unsynchronized.

CAUTION Tuning network timers of any type is an extremely advanced task and should be done only under very special circumstances and with the aid of the Cisco TAC team.

Across WAN networks, it might be advantageous to turn off routing updates completely and to define manually or statically the best path to be taken by the router. Note also that sophisticated routing protocols such as EIGRP or OSPF send out only incremental updates. Be aware, however, that these are correspondingly more complex to design and implement, although ironically, the configuration is very simple.

Another method of reducing routing updates is to implement *snapshot routing*, which is available on Cisco routers and designed for use across on-demand WAN links. This allows the routing tables to be frozen and updated either at periodic intervals or when the on-demand link is brought up. For more information on this topic, refer to the Cisco web page.

To optimize the traffic flow throughout a network, you must carefully design and configure the IP network. In a client/server environment, control of the network overhead is even more important. The following section discusses some concerns and strategies.

Prioritization

Access lists are not used just to determine which packets will be forwarded to a destination. On a slow network connection where bandwidth is at a premium, access lists are used to determine the order in which traffic is scheduled to leave the interface. Unfortunately, some of the packets might time out. Therefore, it is important to carefully plan the prioritization based on your understanding of the network. You need to ensure that the most sensitive traffic (that is, traffic most likely to time out) is handled first.

Many types of prioritization are available. Referred to as *queuing techniques*, they are implemented at the interface level and are applied to the interface queue. The weighted fair queuing (WFQ) technique is turned on by default on interfaces slower than 2 Mbps, and can be tuned with the **fair-queue x y z** interface configuration command.

The WFQ method is available in the later versions of the IOS. It is turned on automatically—in some instances, by the Cisco IOS—replacing the first-in, first-out (FIFO) queuing mechanism as the default. The queuing process analyzes the traffic patterns on the link, based on the size of the packets and the nature of the traffic, to distinguish interactive traffic from file transfers. The queue then transmits traffic based on its conclusions.

Queuing techniques that are manually configured with access lists are as follows:

- **Priority queuing** — This is a method of dividing the outgoing interface buffer into four virtual queues. Importance or priority ranks these queues, and traffic will be sent out of the interface accordingly. This method ensures that sensitive traffic on a slow or congested link is processed first.
- **Custom queuing** — The interface buffer is divided into many subqueues. Each queue has a threshold stating the number of bytes or the number of packets that might be sent before the next queue must be serviced. In this way, it is possible to determine the percentage of bandwidth that each type of traffic is given.
- **Class-based weighted fair queuing (CBWFQ)** — This queuing method extends the standard WFQ functionality to provide support for user-defined traffic classes. For CBWFQ, you define traffic classes based on match criteria, including protocols, access control lists (ACLs)—known as simply access lists in Cisco parlance—and input interfaces. Packets satisfying the match criteria for a class constitute the traffic for that class. A queue is reserved for each class, and traffic belonging to a class is directed to that class's queue.
- **Low-latency queuing (LLQ)** — This feature brings strict priority queuing to CBWFQ. Configured by the **priority** command, strict priority queuing gives delay-sensitive data, such as voice, preferential treatment over other traffic. With this feature, delay-sensitive data is sent first. In the absence of data in the priority queue, other types of traffic can be sent.

Reducing Network Traffic: Alternatives to Access Lists

Because of the resources required to process access lists, they are not always the most suitable solution. The null interface is a good example of when a technology can be used imaginatively to produce a low-resource solution.

The *null interface* is a virtual or logical interface that exists only in the operating system of the router. Traffic can be sent to it, but it disappears because the interface has no physical layer. A virtual interface does not physically exist. Administrators have been extremely creative and have used the

interface as an alternative to access lists. Access lists require CPU processing to determine which packets to forward. The null interface just forwards the traffic to nowhere.

By default, the router responds to traffic sent to the null interface by sending an Internet Control Message Protocol (ICMP) Unreachable message to the source IP address of the datagram. However, you can configure the router simply and silently drop the datagrams. With this configuration, no error messages are sent to the transmitting node. This has several benefits, one of which is additional security.

To disable the sending of ICMP Unreachable messages in response to packets sent to the null interface, in interface configuration mode, type the following:

```
Router(config-if)#no ip unreachable
```

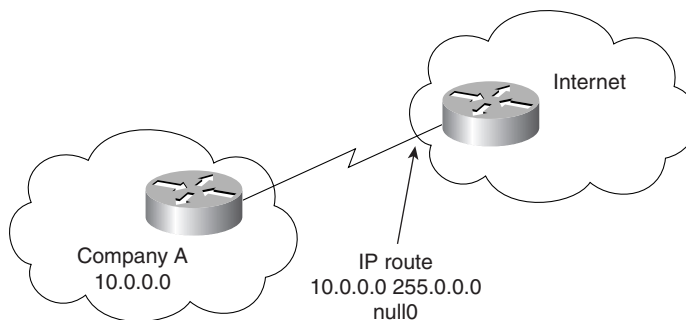
The following sections provide examples of how a null interface can be used within the Internet, as well as in an intranet environment.

Internet Example

If the router receives traffic to be forwarded to network 10.0.0.0, it will be dropped through null0 into a “black hole.” Because this is a private network address to be used solely within an organization, never to stray onto the Internet, this is a command that may well be configured on routers within the Internet.

Figure 3-2 shows how you might implement a null interface in an organization. The example shows how it can be used to filter the private network from entering the Internet.

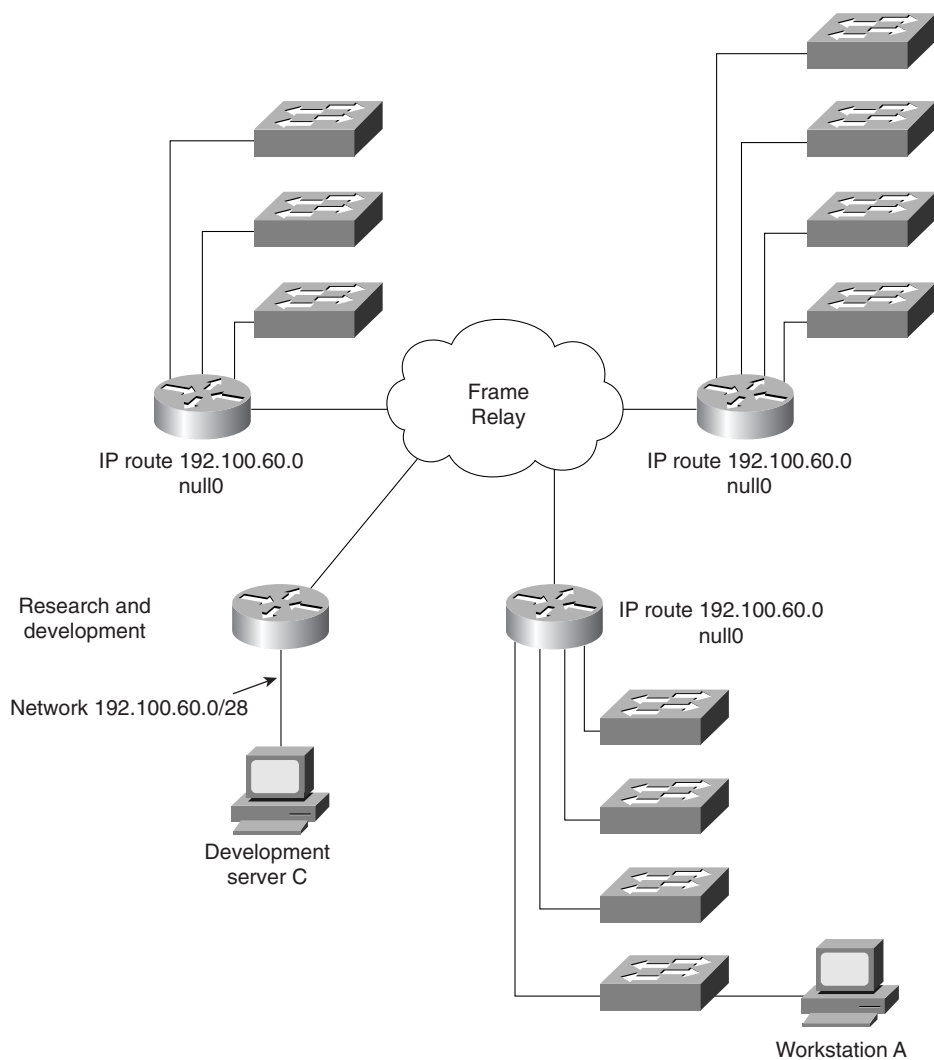
Figure 3-2 *Using the Null Interface on the Internet*



Intranet Example

Configuring the static route to null0 on an internal company router would prevent connectivity to the defined network because all traffic to that destination would be forwarded to the null0 interface and dropped. This is illustrated in Figure 3-3.

Figure 3-3 *Using the Null Interface Within an Organization*



In Figure 3-3, Workstation A would not be capable of connecting to Server C, the development server used by the Research and Development department. The result is that the Research and Development department would be capable of seeing the rest of the organization. Indeed, the rest of

the world can see the Research and Development department in a routing table. Any attempt to direct traffic to the network will be unsuccessful, however. The first router that sees the traffic will statically route it to the null interface, which metaphorically is a black hole.

NOTE Because the static route is entered into the routing table, it is important to remember that all the rules of static routing apply. By default, if the router hears of the destination route via another source, it is ignored in favor of the static route that has a lower administrative distance (more credible source).

Certain guidelines or key points should be used in the design of an IP network. The following section identifies these guidelines.

Keys Points to Remember When Designing an IP Network

When addressing an IP network, you should consider whether it is for an existing network or a network that is to be created from scratch, because the approaches will differ. Because the concerns are different, the following list considers general points that apply to both kinds of network. This is followed by a discussion of points to think about when readdressing an existing network.

You should consider the following list of items when preparing the IP addressing plan for your network, whether it is a new or existing network:

- Identifying how many hosts and subnets will be required in the future requires communication with other departments, in terms of the growth of personnel and the budget for network growth. Without the standard-issue crystal ball, a wider view must be taken at a high level to answer these questions. The answers need to come from a range of sources, including the senior management and executive team of the organization.
- The design of the IP network must take into consideration the network equipment and its vendors. Interoperability may well be an issue, particularly with some of the features offered by each product.
- For route aggregation (summarization) to occur, the address assignments must have topological significance.
- When using VLSM, the routing protocol must send the extended prefix (subnet mask) with the routing update.
- When using VLSM, the routing protocol must do a routing table lookup based on the longest match.

- Make certain that enough bits have been allowed at each level of the hierarchical design to address all devices at that layer. Also be sure that growth of the network at each level has been anticipated. What address space is to be used (Class A, B, C, private, registered), and will it scale with the organization?

NOTE Cisco offers many enhancements in its IOS Software. Most of these enhancements are interoperable. If they are not, Cisco provides solutions for connecting to industry standards (which, of course, are fully supported by Cisco). Check Cisco.com to review the latest features and any connectivity issues.

In many cases, not enough consideration is given to IP address design with regard to the routing process, leaving the decision to be based on the longest address match. Careful consideration of IP addresses is essential to the design of a VLSM network.

Consider a network, as described in Chapter 2 in the section “Assigning IP VLSM Subnets for WAN Connections,” that uses the Class B Internet address 140.100.0.0.

The routing table has the following among its entries:

- 140.100.0.0/16
- 140.100.1.0/20
- 140.100.1.192/26

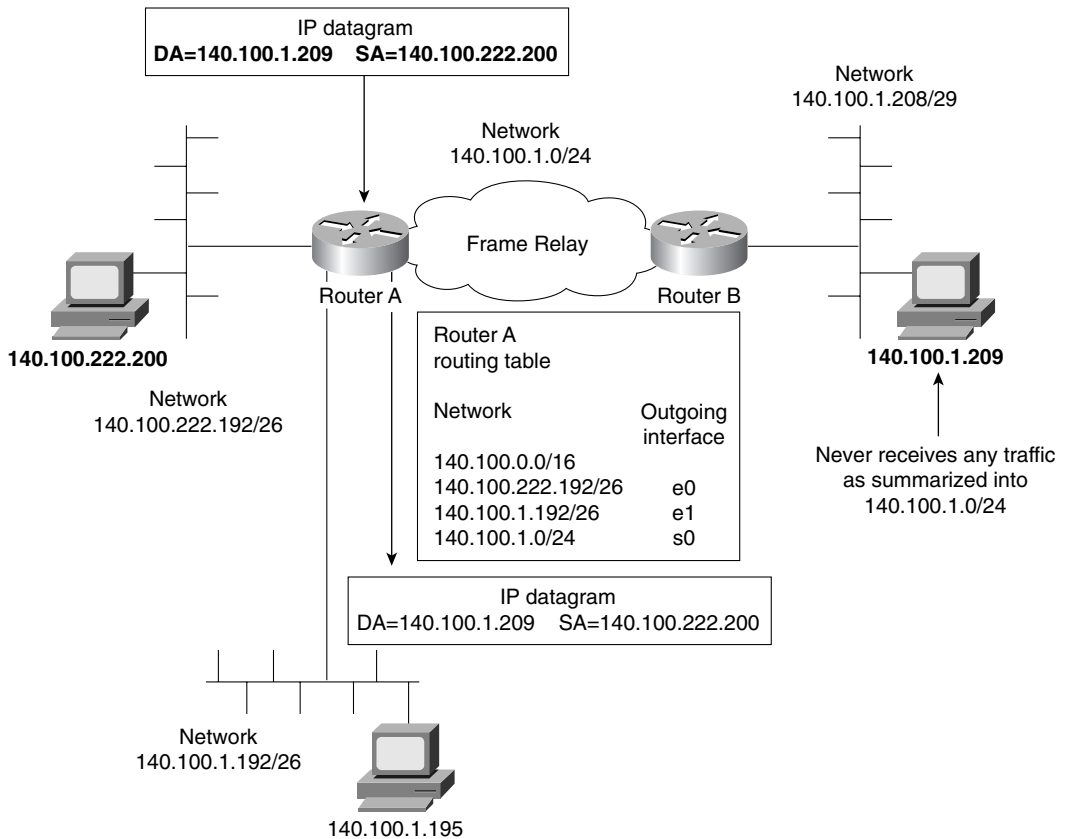
A packet comes into the router destined for the end host 140.100.1.209. The router will forward to the network 140.100.1.192 because the bit pattern matches the longest bit mask provided. The other routes are also valid, however, so the router has made a policy decision that it will always take the most specific mask, sometimes referred to as the *longest match*.

This decision is based on the design assumption that has been made by the router that the longest match is directly connected to the router or that the network is reached from the identified interface. If the end host 140.100.1.209 actually resides on network 140.100.1.208/29, this network must be accessible through the interface that has learned of the subnet 140.100.1.192/26. Summarization will have been configured, because 140.100.1.192 is an aggregate of various networks, including the network 140.100.1.208/29.

If the network 140.100.1.208/29 resides out of the interface that has learned about 140.100.1.0/20, no traffic will ever reach the subnet 140.100.1.208/29, because it will always forward based on the longest match in the routing table. The only solution is to turn off summarization and to list every subnet with the corresponding mask. If summarization is turned off, the subnet 140.100.1.208/29 will not be summarized into the network 140.100.1.0/20. It will consequently be the longest match

in the routing table, and traffic will be sent to the destination network 140.100.1.208/29. Figure 3-4 shows an example of route summarization.

Figure 3-4 *Route Summarization and VLSM*



Designing IP Addresses for an Existing Network

Up to this point, the discussion has dealt with organizations that are designing an IP network for the first time. In reality, this is rarely the case, unless a decision has been made to readdress the entire network.

Often the network has been up and running for some years. If this is the case, the usual task is to use some of the newer technologies available to reduce and manage network traffic so that the network can grow without pain.

The simplest solution is to implement a classless routing protocol that sends the subnet mask in the updates and thus allows VLSM and summarization. OSPF, EIGRP, and IS-IS are examples of

classless routing protocols. For a detailed comparison of the various routing protocols, refer to Chapter 1, “IP Routing Principles,” in the section “Types of Routing Protocols.” However appropriate the routing protocol that you have chosen, it might not be possible to use the summarization feature. As explained earlier, this capability is determined in part by how well the addressing scheme mirrors and is supported by the physical topology.

You can use the following guidelines to determine whether summarization can be configured within a particular network:

- The network addressing scheme should reflect the physical topology of the network.
- The physical and logical topology of the network should be hierarchical in design.
- Given the network addressing scheme, the addresses to be summarized need to share the same high-order bits.
- If the subnet addresses are clearly set on a single binary border, this suggests a prefix mask of /21 or 255.255.248.0. Because the subnets are multiples of 8, they might be summarized by a higher subnet value that is divisible by 8, such as 140.100.64.0. The following subnets provide an example:
 - 140.100.64.0
 - 140.100.72.0
 - 140.100.80.0
 - 140.100.88.0
 - 140.100.96.0
 - 140.100.104.0
 - 140.100.112.0
 - 140.100.120.0
- The nature of the traffic flow within the network should reflect the hierarchical logical and physical design.
- The routing protocol used must support VLSM.

Using this list to identify whether summarization is possible, you might find that you do not have the answers to some of the questions that arise or that another solution to readdressing must be found.

For example, any design of a network requires very careful analysis of the current network and a clear understanding of the organization’s plans. Unfortunately, it is not always possible to determine the nature or flow of data through a network. Intranets and internal web pages have made the nature of the traffic within an organization far more unpredictable.

The increased tendency for organizations to need flexibility or mobility in addressing can make the IP design very challenging. The design would need to include *Dynamic Host Configuration Protocol (DHCP)* and *Domain Name System (DNS)* servers to maximize the flexibility of the network.

DHCP allows end hosts to be assigned an IP address upon application. As an example, consider an airline that assumes that not everyone will turn up for the flight, so it can oversell the seats on the plane. In a similar fashion, the DHCP server has a block of addresses, but it does not expect every machine on the network to turn on at the same time. Thus, 100 users are provided 60 IP addresses.

The DNS server provides a name-to-address translation, which is extremely useful when the DNS server works in conjunction with the DHCP server.

It is also important to understand fully the nature of the traffic in the network, particularly if it is a client/server environment, in which the design must allow for servers to communicate with each other and with their clients.

Using the existing addressing of the organization might not be possible. If this is the case, the decision must be made to readdress the network. You might need to make this decision for two reasons: either the network cannot scale because of the limitations of the classful address that has been acquired from the IANA, or the original design does not allow for the current environment or growth.

If the addressing scheme is inadequate in size, you have several options. The first action for the administrator to take is to apply to the IANA for another address; the second is to use private addressing. The next section describes private addresses on the Internet.

Private Addresses on the Internet

Private addressing is one of the solutions (along with VLSM, IPv6 with an address field of 128 bits, and CIDR addressing and prefix routing) that the Internet community began to implement when it became apparent that there was a severe limitation to the number of IP addresses available on the Internet.

Private addressing is defined by RFC 1597 and revised in RFC 1918. It was designed as an addressing method for organizations that have no intention of ever connecting to the Internet. If Internet connectivity is not required, there is no requirement for a globally unique address from the Internet. The individual organization could address its network without any reference to the Internet, using one of the address ranges provided.

The advantage of the Internet is that none of the routers within the Internet recognize any of the addresses designated as private addresses. If an organization that deployed private addressing as outlined in RFC 1918 (in error) connected to the Internet, all its traffic would be dropped. The

routers of Internet service providers (ISPs) are configured to filter all network routing updates from networks using private addressing. In the past, organizations “invented” addresses, which were, in fact, valid addresses that had already been allocated to another organization. There are many amusing and horrifying stories of organizations connecting to the Internet and creating duplicate addresses within the Internet. A small company inadvertently masquerading as a large state university can cause much consternation.

Table 3-2 outlines the IP address ranges reserved for private addressing, as specified in RFC 1918.

Table 3-2 Private Address Ranges

Address Range	Prefix Mask	Number of Classful Addresses Provided
10.0.0.0 to 10.255.255.255	/8	1 Class A
172.16.0.0 to 172.31.255.255	/12	16 Class Bs
192.168.0.0 to 192.168.255.255	/16	256 Class Cs

The use of private addressing has now become widespread among companies connecting to the Internet. It has become the means by which an organization avoids applying to the IANA for an address. As such, it has dramatically slowed, if not prevented, the exhaustion of IP addresses.

Because private addresses have no global significance, an organization cannot just connect to the Internet. It must first go through a gateway that can form a translation to a valid, globally significant address. This is called a *Network Address Translation (NAT)* or *NAT gateway*.

Configuring private addressing is no more complicated than using a globally significant address that has been obtained from the IANA and is “owned” by the organization. In many ways, configuring private addressing is easier, because there are no longer any restrictions on the subnet allocation, particularly if you choose the Class A address 10.0.0.0.

The reasons for addressing your organization’s network using private addressing include the following:

- There is a shortage of addressing within the organization.
- You require security. Because the network must go through a translation gateway, it will not be visible to the outside world.
- You have an ISP change. If the network is connecting to the Internet through an ISP, the addresses allocated are just on loan or are leased to your organization. If the organization decides to change its ISP, the entire network will have to be readdressed. If the addresses provided define just the external connectivity and not the internal subnets, however, readdressing is limited and highly simplified.

The use of private addressing has been implemented by many organizations and has had a dramatic impact on the design of IP networks and the shortage of globally significant IP addresses. You should bear some things in mind when designing an IP network address plan using private addressing, including the following:

- If connections to the Internet are to be made, hosts wanting to communicate externally will need some form of address translation performed.
- Because private addresses have no global meaning, routing information about private networks will not be propagated on interenterprise links, and packets with private source or destination addresses should be forwarded across such links with extreme care. Routers in networks not using private address space, especially those of ISPs, are expected to be configured to reject (filter out) routing information about private networks.
- In the future, you might be connecting, merging, or in some way incorporating with another company that has also used the same private addressing range.
- Security and IP encryption do not always allow NAT.

If private addressing is deployed in your network and you are connecting to the Internet, you will be using some form of NAT. The following section explains this technology.

Connecting to the Outside World with NAT

When connecting to the outside world, some filtering and address translation might be necessary. Unless an address has been obtained from the Internet or from an ISP, you must perform address translation. The RFC that defines NAT is RFC 1631, “The IP Network Address Translator.”

NAT is the method of translating an address on one network into a different address for another network. It is used when a packet is traversing from one network to another and when the source address on the transmitting network is not legal or valid on the destination network, such as when the source corresponds to a private address. The NAT software process must be run on a Layer 3 device or router (which is logical, because NAT deals with the translation of Layer 3 addresses). NAT is often implemented on a device that operates at higher layers of the OSI model because of their strategic placement in the organization. NAT is often used on a firewall system, for example, which is a security device that guards the entrance into the organization from the outside world. The position of the firewall makes it an excellent choice for NAT, because most translations are required for traffic exiting an organization that has used private addressing as defined in RFC 1918.

NAT had a controversial childhood, particularly when it was used for translating addresses that did not use RFC 1918 guidelines for private addressing; sometimes an organization used an address that had just been created imaginatively by a network administrator. This practice occurred when there was no glimmer of a possibility that the organization would ever connect to the Internet. This

certainty that a company would never connect to the Internet is unrealistic, even for small companies, in an era when even individual homes have Internet connectivity.

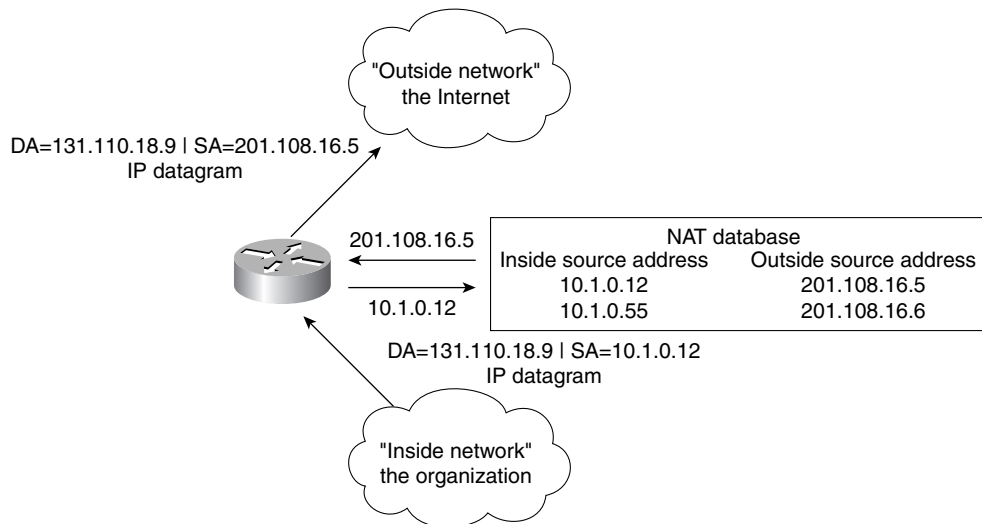
Therefore, NAT is useful in the following circumstances:

- To connect organizations that used address space issued to other organizations to the Internet
- To connect organizations that use private address space defined in RFC 1918 and want to connect to the Internet
- To connect two organizations that have used the same private address, in line with RFC 1918
- When the organization wants to hide its addresses and is using NAT as part of firewall capabilities or is using additional security features

TIP NAT is designed for use between an organization and the outside world. Although it might be used to solve addressing problems within an organization, you should see this as a temporary fix. In such situations, NAT is a transitory solution to keep the network functional while you are designing and readdressing it appropriately.

Figure 3-5 illustrates an organization connecting to the outside world using NAT.

Figure 3-5 *Connecting to the Outside World Using NAT*



Cisco supports the use of NAT on the majority of its platforms, as well as on its Cisco Secure PIX firewall. Various levels of support are offered, depending on the platform and the IOS release that your company has purchased. Cisco now bundles NAT support into the standard product offering. It started to be widely offered from IOS version 11.2 with the purchase of the “plus” software, and full NAT functionality became available in the Base IOS form with version 12.0. NAT itself is currently at version 3.0. The following sections describe the main features and functions of NAT that Cisco offers.

NOTE If you are considering implementing NAT, contact Cisco via its web page. You should always contact the vendor of a product before purchase to appreciate fully the latest offerings and pricing. Because this industry is so dynamic, it is wise to verify the latest data.

The Main Features of NAT

The main features of NAT, as supported by Cisco, include the following:

- **Static addressing** —This one-to-one translation is manually configured.
- **Dynamic source address translation** —Here, a pool of addresses is defined. These addresses are used as the product of the translation. They must be a contiguous block of addresses.
- **Port address translation (PAT)** —Different local addresses (within the organization) are translated into one address that is globally significant for use on the Internet. The additional identifier of a TCP or UDP port unravels the multiple addresses that have been mapped to single addresses. The uniqueness of the different local addresses is ensured by the use of the port number mapped to the single address.
- **Destination address rotary translation** —This is used for traffic entering the organization from the outside. The destination address is matched against an access list, and the destination address is replaced by an address from the rotary pool. This is used only for TCP traffic, unless other translations are in effect.

The Main Functions of NAT

The basic operation of NAT is very straightforward, although the terminology is rather confusing. The list of address definitions in Table 3-3 clarifies the different terms.

To translate one network address into another, the process must differentiate between the functionality of the addresses being translated. Table 3-3 lists the categories of functions.

Table 3-3 Categories of Functions

Address	Definition
Inside Global	The addresses that connect your organization indirectly to the Internet. Typically, these are the addresses provided by the ISP. These addresses are propagated outside the organization. They are globally unique and are the addresses used by the outside world to connect to inside the organization. Simply explained, they are the addresses that define how the <i>inside</i> addresses are seen <i>globally</i> by the outside.
Inside Local	The addresses that allow every end device in the organization to communicate. Although these addresses are unique within the organization, they are probably not globally unique. They may well be private addresses that conform to RFC 1918. They are the <i>inside</i> addresses as seen <i>locally</i> within the organization.
Outside Global	These are the Internet addresses (all the addresses outside the domain of the organization). They are the <i>outside</i> addresses as they appear to the <i>global</i> Internet.
Outside Local	These addresses are external to the organization. This is the destination address used by a host inside the organization connecting to the outside world. This will be the destination address of the packet propagated by the internal host. This is how the <i>outside</i> world is seen <i>locally</i> from inside the organization.

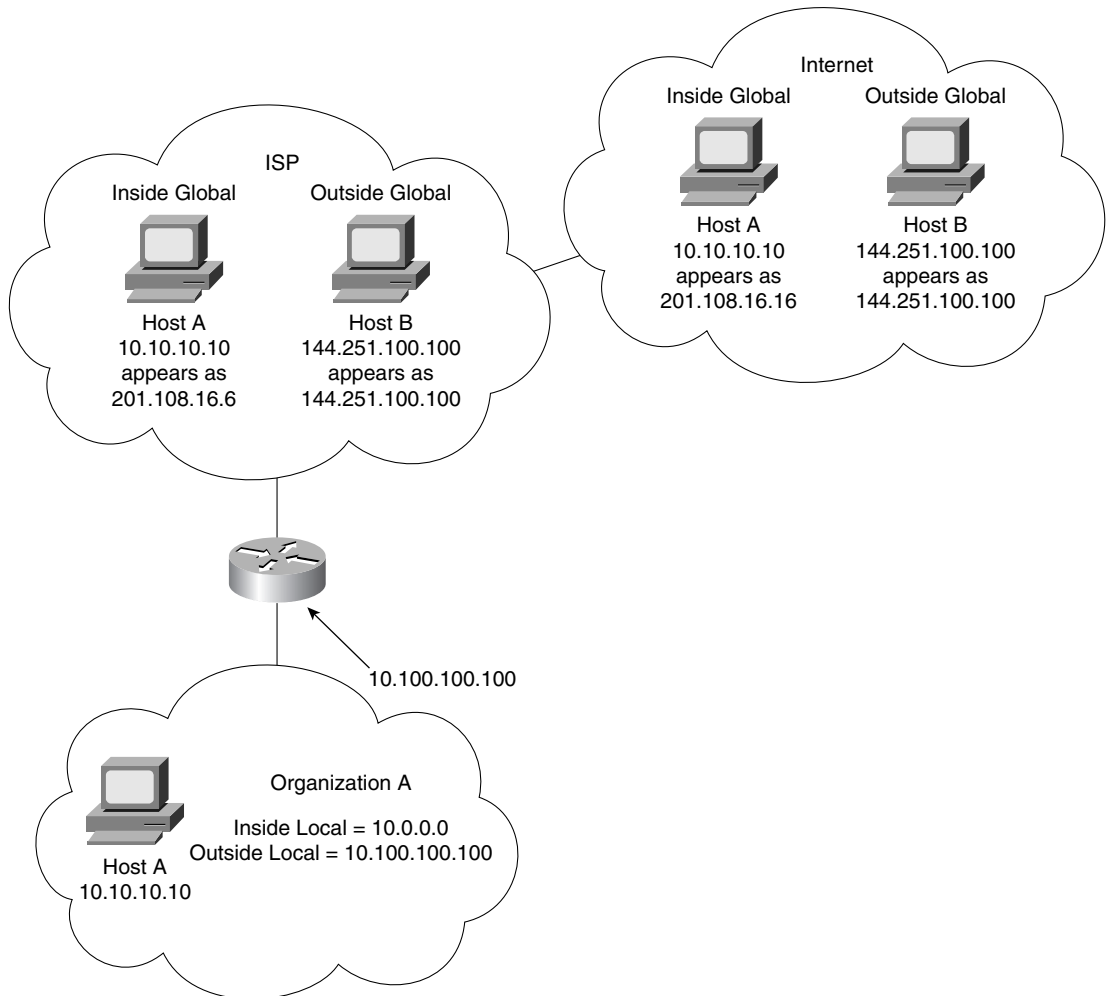
As shown in Figure 3-6, a router within the organization sees the inside addresses and the address of the router connecting them to the outside world, namely the Outside Local address. The router that connects to the outside world has an Inside Global address (how it is seen by the rest of the world) and an address to connect to the ISP, the Outside Global address. The diagram shows what each router sees based on its position in the NAT world.

Figure 3-6 illustrates the terms defined in Table 3-3.

Understanding IPv6

You have learned about IP addressing, but the discussion so far has been about IPv4, how to address a network, and how to overcome some of its limitations. IPv6 is the solution to many of the limitations in addressing that are seen in IPv4. Although there are IPv6 implementations, IPv6 is yet to be seen as a widespread solution, mainly because of the overwhelming task of readdressing networks and upgrading applications. Both NAT and private addressing are creative solutions to the inherent capacity problem that IPv4 has encountered. As the demand for IP addresses increases, these solutions, however creative, cease to be effective.

IPv6 quadruples the address space, providing 128 bits instead of the 32 bits currently available with IPv4. In real terms, that increases the number of addresses from just more than four million to a nearly infinite number of addresses. The address size is quadrupled, allowing approximately 1030 addresses per person on the planet.

Figure 3-6 *Using the NAT Terms*

With IPv6, the ability to dispense with solutions such as NAT, private addresses, and temporarily assigned addresses through DHCP means that end-to-end connectivity is available. With this direct connectivity come some technical enhancements. Both security and QoS might be implemented more efficiently when there is end-to-end connectivity, with no intermediary translations.

IPv6 offers the following benefits and features:

- Larger address space
- Unicast and multicast addressing
- Address aggregation

- Autoconfiguration
- Renumbering
- A simple and efficient header
- Security
- Mobility
- Options for transitioning from IPv4 to IPv6
- Routing protocols

The following sections describe each of these features in detail.

IPv6 Address Format

The IPv6 address is very different from the IPv4 address. Not only is it four times the length, increasing the length from 32 to 128 bits, but it is also represented in hexadecimal as opposed to decimal notation. Colons separate the 16-bit hexadecimal number fields, which are portions of the 128-bit address, from the 128 bits.

An example of an IPv6 address follows:

4021:0000:240E:0000:0000:0AC0:3428:121C

To avoid confusion, error, and unnecessary complication, the following rules have been determined. These rules simplify the address where possible, making it more manageable:

- The hexadecimal numbers are not case sensitive, preventing operator error in entering addresses.
- Leading 0s in any 16-bit field can be dropped and represented by colons.
- A pair of colons (::) indicates the successive 16-bit fields of 0s have been dropped. The process easily identifies the number of 0s dropped by adding 0s until the address is once again 128 bits long.
- Only one pair of colons is allowed in any address, because the process would not be able to identify how many 0s should be replaced in each location.

NOTE The rules for the addressing of IPv6, including guidelines for simplification, are given in the RFC 2373, “IP Version 6 Addressing Structure.”

Keeping these rules in mind, the following address:

4021:0000:240E:0000:0000:0AC0:3428:121C

can be written in the following form:

4021:0:240E::0AC0:3428:121C

Although there cannot be two instances of a double colon, those fields with only 0s can be shown as 0. In this example, the second field shows the 0s reduced to one representational zero.

If the address is that of a network with no host ID shown, the address can end in a double colon, for example:

4021:0:240E::

IPv6 addressing comes in many forms, and it is able to solve many of the limitations of IPv4 not simply through additional bits but through greater flexibility and complexity.

IPv6 Unicast Addresses

The IPv6 unicast addresses are divided up according to functionality. Although a unicast address is tied to a specific node with a unique address to identify it, the scope of the search for that end system is clearly defined in IPv6. This minimizes the resources required, making the transport of packets across the network faster and more efficient all around.

IPv6 unicasts come in the following flavors:

- **Link local** —This is a specific address, known as a local link unicast address, where the end system is on the same physical link. This would include discovery protocols, routing protocols, and other control protocols. These addresses are autoconfigured and use the prefix FE80::/10.
- **Site local** —This is a system that is within the same site but might be on a different network. It requires no connection to the global network of the Internet, because there is no guarantee that the addressing is globally unique.
- **Aggregate global unicast** —This is an Internet address that is globally unique.
- **Unspecified and loopback** —This address is simply a placeholder, often used when downloading software or requesting an address. The loopback address is used to test the interface in basic troubleshooting. The address is:

0000.0000.0000.0000.0000.0000.0000.0000.0000.0000.0000.0001

or

0.0.0.0.0.0.0.0.0.0.1

or

::1

IPv6 Multicast Addresses

A multicast address is an address that identifies a group of interfaces, typically on different end systems. The packet is delivered to all the systems identified in the multicast address.

Using multicast addresses is much more efficient than using broadcasts, which require every end system to stop what it is doing, taking both time and resources. Because a multicast address is an address to a group of systems, if the receiving system is not part of the multicast group, it discards the packet at Level 2. However, broadcasts are processed through the OSI stack before the system can determine that the broadcast is not relevant to them.

Layer 2 devices (bridges and switches) propagate broadcasts because broadcast addresses are not stored in its forwarding CAM table. Unlike a router, whose default is to drop packets with unknown addresses, a switch will propagate a frame with an unknown destination address out of every interface. Theoretically, this is also true of multicast addresses, though some devices have intelligence built into the software to restrict multicast propagation. The LAN technologies can propagate these broadcasts around and around if there is a problem, thus causing a broadcast storm that can seriously affect response time and, in extreme cases, network connectivity.

IPv6 does not use broadcasts at all, relying solely on the use of multicast addresses. Though IPv4 uses multicasts as defined in RFC 2365, “Administratively Scoped IP Multicast,” it uses them in a different manner. The IPv6 multicast has a much larger address range.

All IPv6 multicast addresses start with the first 8 bits of the address set to 1. Thus all multicast addresses start with the hexadecimal notation FF (1111 1111). The multicast range is as follows:

FF00::/8
FFFF::/8

The second octet, following the first octet of FF, identifies both the scope and the lifetime of the multicast address. In this way, IPv6 has millions of group multicast addresses to use in current and emerging technologies.

Address Aggregation

Summarization, wherever possible, is crucial within the Internet. The current offering of IPv4 and the routing tables makes summarization critical. The routing tables are more manageable with the implementation of CIDR. Although the addressing scheme in IPv6 allows for an almost infinite amount of addresses to be allocated, the address structure must employ a hierarchical structure so as not to overrun itself.

As in IPv4, the leftmost bits of the address are used to summarize networks that appear lower in the bit structure. Thus, the IPv4 address 140.108.128.0/17 could include the subnets 140.108.128.0/24 through to 140.108.255.0/24. This would mean that the routing tables could route to all the subnets,

but that instead of having 128 subnets listed in the routing tables, there is a single entry. To locate a minor subnet, the normal rules of routing are followed and the packet is sent to the router advertising 140.108.128.0/17. This router, armed with the more detailed routing table, forwards the packet on until it reaches the destination.

In IPv6, the address structure allows for greater granularity in the external address format used within the Internet. The address is very long, and each part serves a function. The first 48 bits of the address are a header used by the IANA for external routing within the Internet to create the Aggregate Global Unicast. The first 3 bits (or the 3 leftmost bits) are fixed as 001 to indicate a global address.

The Site Level Aggregator (SLA) is the address used for routing within the autonomous system and identifies the destination network. It can be used without the 48-bit prefix assigned by the IANA. If this 48-bit prefix is not granted or used, the addressing scheme is similar to IPv4 private addressing, and the autonomous system must not attach to the Internet.

The interface address is often autoconfigured by using the MAC address of the interface.

The IPv6 address that is unique to the Internet is called the Aggregate Global Unicast. The various components described are summarized to the bit allocation below, showing the following address structure:

A fixed prefix of 001	3 bits
IANA allocated prefix	45 bits
Site Level Aggregator (SLA)	16 bits
Interface	64 bits

Autoconfiguration

The local or directly connected router sends out the prefix of the local link and the router's default route. This is sent to all the nodes on the wire, allowing them to autoconfigure their own IPv6 addresses. The local router provides the 48-bit global prefix and the SLA or subnet information to each end system. The end system simply adds its own Layer 2 address, which is unique because it is the burned-in address (MAC address) taken from the interface card. This MAC address, when appended to the 48-bit global address and the 16-bit subnet address, makes up the unique 128-bit IPv6 end system address.

The ability to simply plug in a device without any configuration or DHCP server allows new devices to be added to the Internet, such as cell phones, wireless devices, and even the home toaster. The Internet has become "plug-and-play."

Renumbering

The ability to connect remote devices automatically alleviates many other tasks that were previously administrative nightmares, requiring months of project planning. In IPv4, the mere thought of readdressing the network made experienced, competent network managers turn pale and shake. IPv6 autoconfiguration allows the router to provide the required information to all the hosts on its network. This means they can renumber or reconfigure their address with ease. This is a requirement if and when you change service providers, because the service provider issues the addressing scheme for its customers. With IPv6, it is reassuring to know that such a radical change can be transparent to the end user.

Simple and Efficient Header

The IPv6 header has been simplified to speed up processing and, thus, the performance and efficiency of the router. This has been achieved in the following ways:

- There are fewer fields in the header.
- The fields are aligned to 64 bits.
- The checksum is removed.

The reduction in processing is because of the fewer fields to process. Memory is used more efficiently with the fields aligned to 64 bits. This allows the lookups to be very fast, because the 64-bit fields take advantage of the 64-bit processors in use today. The only drawback is the use of the 128-bit address, which is larger than the current atomic word size.

The removal of the checksum reduces the processing time further. A calculation has been moved up the stack to the transport layer, where both the connection and connectionless transport are required to issue checksums. Remember that the improved efficiency is realized at each router in the path to the destination host, which greatly increases the overall efficiency.

Figure 3-7 compares the IPv4 header with the IPv6 header.

IPv6 Extension Header

Instead of including the Options field within the header as IPv4 does, IPv6 attaches the Options field to the end of the header, indicating with the Next Header field whether there is something additional to process. This speeds up the processing and also allows for protocol evolution, because many extension fields can be chained together.

Figure 3-7 *The IPv4 and IPv6 Headers Compared*

Version	IHL	ToS	Total Length	
Identification			Flags	Fragment Offset
Time to Live		Protocol	Header Checksum	
Source Address				
Destination Address				
Options			Padding	

Version	Traffic Class	Flow Label	
Payload Length		Next Header	Hop Ltd
Source Address			
Destination Address			

-  Name Change but Same Functionality in IPv6
-  No Change in IPv6
-  Removed in IPv6
-  New to IPv6

Security

With direct end-to-end connectivity achieved with a larger address space, security is a more realistic option with IPv6. Because the need for firewalls and NAT processes between the end hosts is decreased, a wider and more direct approach can be given to security by placing the encryption within the host systems.

Although IPSec is available with IPv4, it is mandatory in IPv6. The use of extension headers allows for a protocol dedicated to end-to-end security.

Mobility

IPv6 was designed with mobility built into the protocol using Mobile IP, which is an Internet Engineering Task Force (IETF) standard. Mobile IP allows end systems to move location without the connection being lost, which is essential for wireless products, such as IP phones and GPS systems in cars.

The IPv6 routing header allows the end system to change its IP address by using a home address as the source of the packets. The home address is stable, allowing the roving address to maintain mobility.

Although IPv4 offers Mobile IP, it does so by tunneling back to the home network and then forwarding the data to the final destination. This is called *triangle routing*, and though it works, it is more cumbersome than the solution that is provided by IPv6.

IPv4 to IPv6 Transitions

The key to the success of IPv6 lies not only in its functionality and efficiency as a routed protocol, but also in the ability to transition existing networks to the new protocol. This requires many things to happen, including the following:

- New addressing
- The installation of a new protocol stack
- New applications that can communicate with the new stack

The main theory is that you should start by deploying IPv6 at the outer edges of the network and move into the core of the network in a slow, methodical, and controlled manner. This means that one of three options must occur: The IPv6 traffic needs to be carried through the IPv4 network so that IPv6 can communicate with other devices in a remote domain; both IPv4 and IPv6 need to run through the network, allowing both protocols to live in peaceful coexistence; or one protocol needs to be translated into the other.

The following methods describe how a transition from IPv4 to IPv6 could occur:

- **IOS dual stack** —Both IPv4 and IPv6 run on all systems. This approach allows new IPv6 applications to be introduced on the end systems. The application on the end system requests either an IPv4 address or an IPv6 address from the DNS server. This determines which application uses which IP protocol.
- **Configured tunnels** —These tunnels are for more permanent solutions and provide a secure and stable method of communication across an IPv4 backbone. Both end points of the tunnel need to be manually configured and to be running both IPv4 and IPv6.
- **6to4 tunneling** —This allows IPv6 to be run over an automatically configured tunnel. It requires that the routers connecting the IPv6 remote sites through the IPv4 cloud need to be running dual stacks. The edge routers responsible for running the tunnel will use the prefix 2002::/16 and append the IPv4 interface address to create an address. The interface is the IPv4 address converted to hexadecimal and added to the routing prefix 2002::/16.

IPv6 Routing Protocols

The IP routing protocols that support IPv6 are RIPng, OSPF, IS-IS, and BGP-4, as of Cisco IOS Software Release 12.2T and later.

RIPng is an interior routing protocol and is supported by Cisco IOS. Its functionality is that of RIPv2. It is a distance vector routing protocol, which means that it uses split horizon with poison reverse and has a maximum hop count. You will learn more about distance vector protocols in Chapter 4, “IP Distance Vector Routing Principles.”

To accommodate the needs of IPv6, RIPng incorporates the following features:

- A multicast address for routing updates
- IPv6 prefixes
- IPv6 as the encapsulation packet for the routing updates

BGP-4+ is an exterior routing protocol. It is used to connect autonomous systems across the Internet or within organizations.

Foundation Summary

The “Foundation Summary” section of each chapter lists the most important facts from the chapter. Although this section does not list every fact from the chapter that will be on your exam, a well-prepared candidate should, at a minimum, know all the details in each “Foundation Summary” before going to take the exam.

The three layers of a hierarchy are as follows:

- The access layer
- The distribution layer
- The core layer

Queuing techniques that are manually configured with access lists are as follows:

- Priority queuing
- Custom queuing
- Class-based weighted fair queuing (CBWFQ)
- Low-latency queuing (LLQ)

Table 3-4 outlines the IP address ranges reserved for private addressing, as specified in RFC 1918.

Table 3-4 *Private Address Ranges*

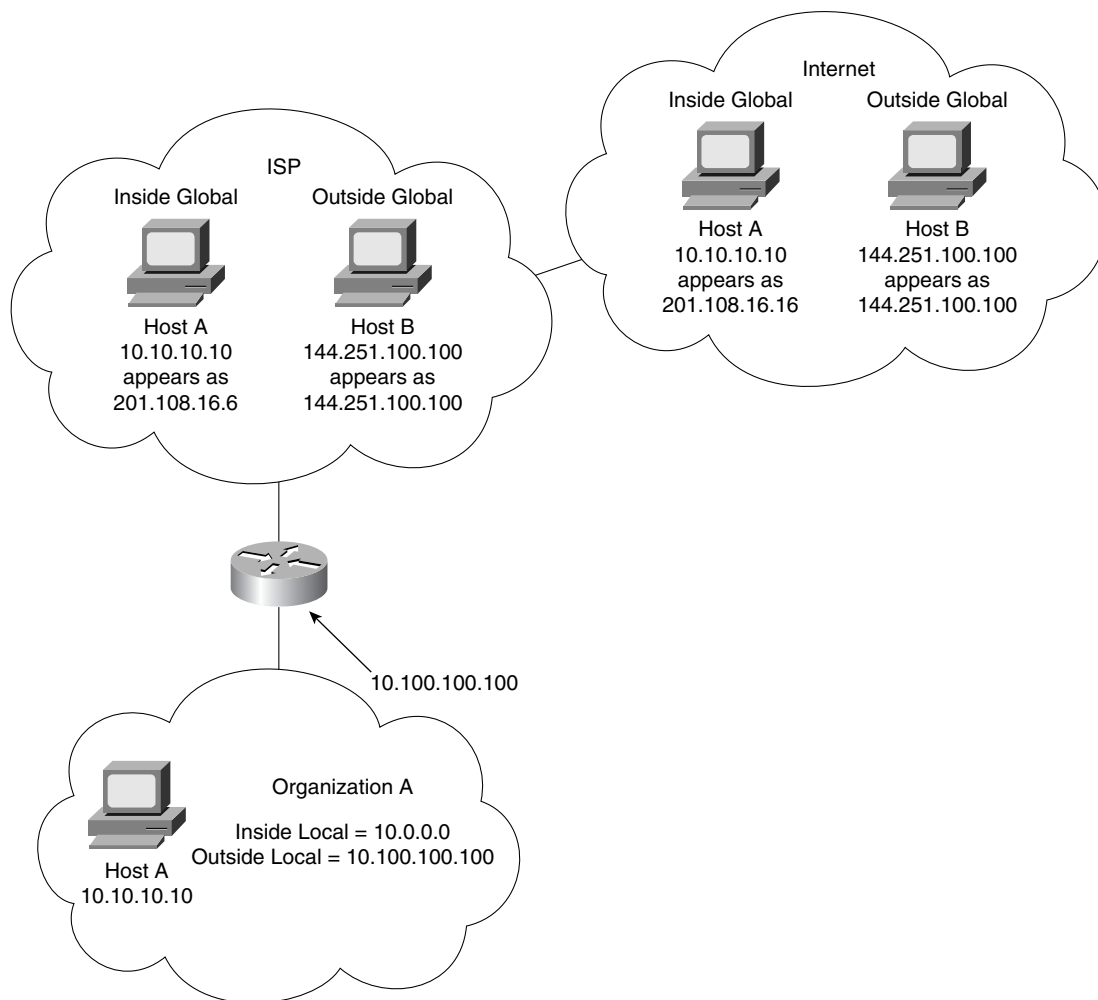
Address Range	Prefix Mask	Number of Classful Addresses Provided
10.0.0.0 to 10.255.255.255	/8	1 Class A
172.16.0.0 to 172.31.255.255	/12	16 Class Bs
192.168.0.0 to 192.168.255.255	/16	256 Class Cs

To translate one network address into another, the process must differentiate between the functionality of the addresses being translated. Table 3-5 lists the categories of functions.

Table 3-5 *Categories of Functions*

Address	Definition
Inside Global	These are the addresses that connect your organization indirectly to the Internet. Typically, these are the addresses provided by the ISP. These addresses are propagated outside the organization. They are globally unique and are the addresses used by the outside world to connect to inside the organization. Simply explained, they are the addresses that define how the <i>inside</i> addresses are seen <i>globally</i> by the outside.
Inside Local	These are the addresses that allow every end device in the organization to communicate. Although these addresses are unique within the organization, they are probably not globally unique. They may well be private addresses that conform to RFC 1918. They are the <i>inside</i> addresses as seen <i>locally</i> within the organization.
Outside Global	These are the Internet addresses (all the addresses outside the domain of the organization). They are the <i>outside</i> addresses as they appear to the <i>global</i> Internet.
Outside Local	These addresses are external to the organization. This is the destination address used by a host inside the organization connecting to the outside world. This will be the destination address of the packet propagated by the internal host. This is how the <i>outside</i> world is seen <i>locally</i> from inside the organization.

Figure 3-8 illustrates the use of NAT terms.

Figure 3-8 *Using the NAT Terms*

IPv6 provides the following features to allow IP networks to scale in a way that IPv4 could not:

- **128-bit address** — The increased address space is a fundamental feature of IPv6. The address has been increased from 32 bits in IPv4 to 128 bits in IPv6.
- **The new header** — The new format increases efficiency in routing. The 64-bit alignment of the fields means the packets are processed at higher speeds. Unnecessary fields have been removed, further streamlining the routing process. A new extension header has been added for optional fields.

- **Autoconfiguration** — This eliminates the need for DHCP servers and manual IP addressing, thus easing network administration and reducing the volume of network errors due to misconfiguration. Not only is end system address acquisition automated, but it also allows for the network to be renumbered or readdressed without visiting each end system for reconfiguration.
- **Security and mobility** — These are built into the protocol specification, as opposed to being configurable options. Both security and mobility are enhanced by the ability to have end-to-end connectivity because of the greater address space available.
- **Transitioning from an IPv4 network** — Transitioning an organization is still complex, requiring much consideration. However, transitioning schemes have been carefully thought through and integrated into the protocol functionality. Two of the most common methods include:
 - Dual stack
 - 6to4 or manually configured tunneling

IPv6 offers the following benefits and features:

- Larger address space, allowing for a larger number of systems that can be globally addressed and a more scalable network
- Increased address space, allowing for a deeper hierarchical structure
- Simplified header, allowing for greater routing efficiency and thus network performance
- Policies for network architecture flexibility, allowing evolution and growth of the protocol
- Support for routing and route aggregation
- Simple administration through serverless autoconfiguration, the ability to renumber with ease, and multihoming, all of which allow a level of plug-and-play support
- Security using IP Security (IPSec) support for all IPv6 devices
- Support for Mobile IP and mobile computing devices (direct-path)
- Multicast support built into the protocol using a greater number of addresses and efficient mechanisms

Q&A

As mentioned in the introduction, “All About the CCNP, CCDP, and CCIP Certifications,” you have two choices for review questions. The questions that follow next give you a bigger challenge than the exam itself by using an open-ended question format. By reviewing now with this more difficult question format, you can exercise your memory better and prove your conceptual and factual knowledge of this chapter. The answers to these questions are found in Appendix A.

For more practice with examlike question formats, including questions using a router simulator and multichoice questions, use the exam engine on the CD.

1. What is an Inside Global address, and when is it used?
2. When is the Inside Local address used?
3. Explain the difference between the Outside Global address and the Outside Local address.
4. Give one example of when NAT might be employed.
5. Explain what PAT is.
6. Why is NAT often configured on the organization’s firewall?
7. Give one example of when private addressing would be a good solution for an organization.
8. Why does summarization need a hierarchical addressing structure?
9. When might you use Destination Address Rotary Translation?
10. Why do NAT and private addressing tend to be implemented together?
11. Which routing protocols for IPv6 does Cisco IOS support?
12. How many bits are there in the IPv6 address space?
13. Explain why it is important to remember to identify how many hosts and subnets are required when designing an IPv4 network.
14. Why does IPv6 addressing allow for more effective security and QoS to be implemented?
15. Give one reason it might be advisable to implement private addressing.
16. In the following address
4021:0000:240E:0000:0AC0:3428:121C
which part of the address is autoconfigured?
17. Why is IPv6 multicasting more efficient than IPv4 broadcasting?
18. State one of the main benefits of IPv6.
19. What is an IPv6 extension header?
20. What are the two most common methods of transitioning an IPv4 network to an IPv6 network?

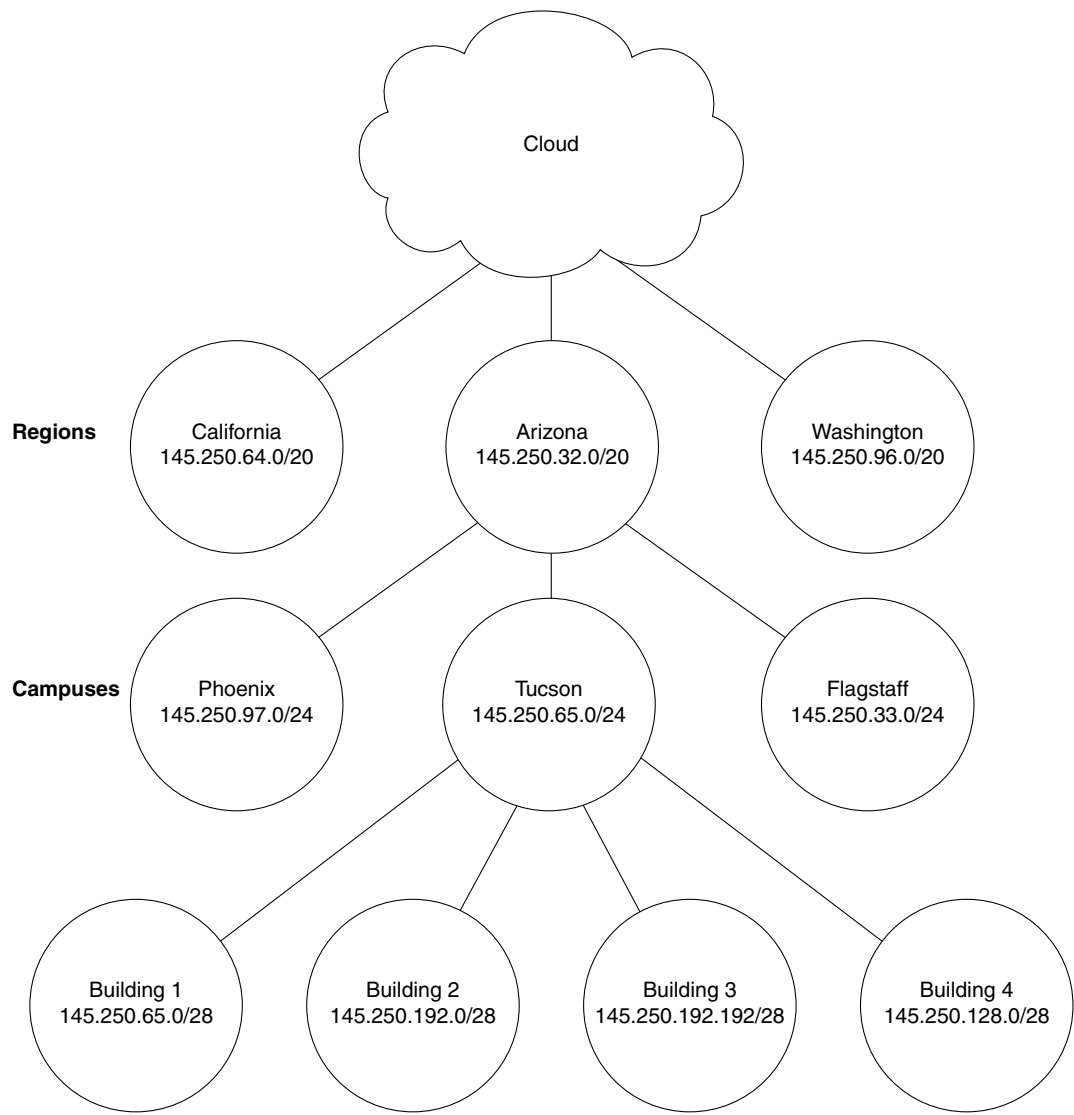
Scenarios

The following scenario and questions are designed to draw together the content of the chapter and exercise your understanding of the concepts. There is not necessarily a right answer. The thought process and practice in manipulating the concepts is the goal of this section. The answers to the scenario questions are found at the end of this chapter.

Scenario 3-1

This scenario concentrates on correcting an addressing scheme in the company discussed in Chapter 2, CyberKit. A network administrator devised the addressing scheme before there was any intention of connecting the company to the Internet and before the company had regional offices. Addresses were subsequently allocated without any policy or administrative control. This has led to problems in the current organization, which now needs to summarize its addresses. Using the addressing scheme in Figure 3-9, answer the following questions.

Figure 3-9 An Addressing Scheme for Scenario 3-1



1. There are serious problems with the addressing scheme in Figure 3-9. If the network had this addressing scheme, would summarization be possible?
2. Design an alternative addressing scheme using VLSM that would summarize to the regional level.
3. Write out the addressing scheme in both binary and dotted decimal notation.

4. Could these addressing requirements be achieved with a Class C address?
5. If the answer to the preceding question is yes, write out the dotted decimal and binary notation to support it. If the answer is no, how many Class C addresses would be required? (Again, write out the dotted decimal and binary notation to support your argument.)

Scenario Answers

The answers provided in this section are not necessarily the only possible answers to the questions. The questions are designed to test your knowledge and to give practical exercise in certain key areas. This section is intended to test and exercise skills and concepts detailed in the body of this chapter.

If your answer is different, ask yourself whether it follows the tenets explained in the answers provided. Your answer is correct not if it matches the solution provided in the book, but rather if it has included the principles of design laid out in the chapter.

In this way, the testing provided in these scenarios is deeper: It examines not only your knowledge, but also your understanding and capability to apply that knowledge to problems.

If you do not get the correct answer, refer back to the text and review the subject tested. Be certain to also review your notes on the question to ensure that you understand the principles of the subject.

Scenario 3-1 Answers

1. *There are serious problems with the addressing scheme in Figure 3-9. If the network had this addressing scheme, would summarization be possible?*

Summarization is not possible, for the following reasons:

- The buildings do not share the same high-order bits as the campus.
- The campuses do not share the same high-order bits as the region.
- Depending on the physical design, the California campus and Building 1 could be seen as duplicate addresses.

2. *Design an alternative addressing scheme using VLSM that would summarize to the regional level.*

See Table 3-6.

3. *Write out the addressing scheme in both binary and dotted decimal notation.*

See Table 3-6.

Table 3-6 shows an alternative solution using the same address as before and the same bit allocation. If you have changed the bit allocation, ensure that there are enough bits for each level of the network. The requirements have not been stated, so you will have to state them for

yourself or use the limited information that is provided. The question identifies three states, and you can assume three campuses in each state. Each campus has four buildings.

Table 3-6 *Alternative Addressing Scheme*

Entire Address in Decimal	Third and Fourth Octets in Binary	Prefix	Subnets	Hosts
145.250.16.0	00010000.00000000	/20	14	4094
145.250.32.0	00100000.00000000	/20	14	4094
145.250.48.0	00110000.00000000	/20	14	4094
145.250.17.0	00010001.00000000	/24	254	254
145.250.18.0	00010010.00000000	/24	254	254
145.250.19.0	00010011.00000000	/24	254	254
145.250.19.16	00010011.00010000	/28	4094	14
145.250.19.32	00010011.00100000	/28	4094	14
145.250.19.48	00010011.00110000	/28	4094	14
145.250.19.64	00010011.01000000	/28	4094	14

4. *Could these addressing requirements be achieved with a Class C address?*

It would not be possible to address this network, using a hierarchical design, with one Class C address. Given the minimum requirements shown in the question of three states, three campuses, and four buildings at each campus, 6 bits would be required. A Class C address allows only 8 bits in total, leaving 2 bits for host allocation. The rule of not using all 0s or all 1s applies to the host portion of the address, so 2 bits would not enable you to address only two hosts.

5. *If the answer to the preceding question is yes, write out the dotted decimal and binary notation to support it. If the answer is no, how many Class C addresses would be required (again write out the dotted decimal and binary notation to support your argument)?*

Remember that although a Class C address has a prefix of /24, not all addresses with a /24 prefix are Class C addresses. A Class C address is a classful address that has been obtained from the IANA. It is interesting that although the first guess is that two are better than one, two Class C addresses do not really improve the situation. The need to address 12 buildings requires 4 bits,

which would allow only 14 hosts in each building. The network could be addressed with two Class C addresses if 14 hosts in each building are all that is required. There is very little growth allowance in this scheme, making it inadvisable.

The most efficient addressing scheme with Class C addresses would be to use 40 Class C addresses. Consider, for example, the addressing scheme using Class C addresses.

A Class C address would be allocated to each building. This would allow 254 hosts in each building and subnetting to the floor, if necessary. The other three Class C addresses would be used with VLSM to identify the regions and campuses. Table 3-7 shows the addressing scheme for the one Class C address to address one region, three campuses, and four buildings.

Each region or state will now advertise five networks—the four Class C addresses for the buildings and the shared network for the state. One Class C network can be used for the state if the connections are point to point. Because there are 15 connections—four buildings per region, and three regions—this means 12 connections to the buildings, plus three connections to the state. A Class C address would easily accommodate this, even with redundant connections built into the design.

The reason that 40 Class C networks are needed is that the analysis of the state must be extrapolated to the entire organization. The organization covers three states, each with three regions, and each region has four buildings. Although the addressing described previously is correct, it would need to be extended to the other regions. This is calculated as follows:

The number of buildings requiring Class C networks in three states, each with three regions, and each region in turn with four buildings is $3 * 3 * 4 = 36$. Add to the three states requiring Class C networks the additional network required for the core cloud that connects the states, and you have $36 + 3 + 1 = 40$. In total, therefore, 40 Class C networks will be required.

Other than for academic interest in torturous addressing, this scenario would be an excellent candidate for a private Class B address.

Table 3-7 *The Class C Used to Identify the Campuses for Each Region*

Entire Address in Decimal	Fourth Octet in Binary	Prefix Mask
Region: Arizona		
210.10.32.0		
Campuses		
210.10.32.32	001 00000	/27
210.10.32.64	010 00000	/27
210.10.32.96	011 00000	/27
Buildings		
Tucson		
210.10.32.36	001001 00	/30
210.10.32.40	001010 00	/30
210.10.32.44	001011 00	/30
210.10.32.48	001100 00	/30
Flagstaff		
210.10.32.68	010001 00	/30
210.10.32.72	010010 00	/30
210.10.32.76	010011 00	/30
210.10.32.80	010100 00	/30
Phoenix		
210.10.32.100	011001 00	/30
210.10.32.104	011010 00	/30
210.10.32.108	011011 00	/30
210.10.32.112	011100 00	/30

Table 3-8 shows how to address the departments or floors within each building. For this discussion, use 210.10.64.0 as the example Class C address. Four bits taken in the fourth octet allows 14 networks, either distributed between the floors or between departments, with 14 hosts on each subnet.

Table 3-8 *How to Address a Building Using a Class C Network Address*

Entire Address in Decimal	Fourth Octet in Binary	Prefix Mask
210.10.64.16	0001 0000	/28
210.10.64.32	0010 0000	/28
210.10.64.48	0011 0000	/28
210.10.64.64	0100 0000	/28
210.10.64.80	0101 0000	/28
210.10.64.96	0110 0000	/28
210.10.64.112	0111 0000	/28
210.10.64.128	1000 0000	/28
210.10.64.144	1001 0000	/28
210.10.64.160	1010 0000	/28
210.10.64.176	1011 0000	/28
210.10.64.192	1100 0000	/28
210.10.64.208	1101 0000	/28
210.10.64.224	1110 0000	/28



This chapter covers the following topics, which you need to understand to pass the CCNP/CCDP/CCIP BSCI exam:

- Distance vector routing protocols
- RIP version 1 and RIP version 2
- IGRP and EIGRP
- Choosing between routing protocols using administrative distance
- Convergence
- Interior and exterior gateway protocols

IP Distance Vector Routing Principles

The topics in this chapter deal with the original routing protocols and the principles that underlie the technology. These routing protocols are the basis of many subsequent subjects tested in the BSCI exam.

In this chapter, the concepts of routing with IP and the mechanics of the process are dealt with in the context of distance vector routing protocols. The topics will directly reflect questions on the BSCI exam. If you do not understand the contents of this chapter, it will be difficult to follow some of the subsequent subjects, making it difficult to pass the exam.

“Do I Know This Already?” Quiz

The purpose of the “Do I Know This Already?” quiz is to help you decide what parts of this chapter to use. If you already intend to read the entire chapter, you do not necessarily need to answer these questions now.

The 18-question quiz, derived from the major sections in “Foundation Topics” portion of the chapter, helps you to determine how to spend your limited study time.

Table 4-1 outlines the major topics discussed in this chapter and the “Do I Know This Already?” quiz questions that correspond to those topics.

Table 4-1 “Do I Know This Already?” Foundation Topics Section-to-Question Mapping

Foundation Topics Section	Questions Covered in This Section
Distance Vector Routing Protocols	1–3
RIPv1 and RIPv2	4–5
IGRP and EIGRP	6–9
Choosing Between Routing Protocols Using Administrative Distance	10–12
Convergence	13–15
Interior and Exterior Gateway Protocols	16–18

CAUTION The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark this question wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

1. Which of the following techniques are used by a distance vector routing protocol?
 - a. Count to infinity
 - b. Poison reverse
 - c. Split horizon
 - d. Feasible successor
2. Which algorithm is used by distance vector routing protocols?
 - a. DUAL
 - b. Bellman Ford
 - c. Dijkstra
 - d. Feasible successor
3. Which of the following are distance vector routing protocols?
 - a. EIGRP,BGP-4
 - b. IGRP, OSPF, IS-IS
 - c. RIPv1, RIPv2, IGRP
 - d. IS-IS, RIPv2
4. Name the interior routing protocols on a Cisco router that sends a routing update every 30 seconds by default.
 - a. IGRP
 - b. RIPv1
 - c. RIPv1, RIPv2
 - d. BGP-4
5. Which of the following are enhancements in the creation of RIP version 2?
 - a. Using the DUAL algorithm
 - b. Maintaining a topology map

- c. Multicast addressing
 - d. Sending the mask with the update
6. How often does IGRP send a routing update (by default)?
- a. 30 seconds
 - b. 90 seconds
 - c. It only sends incremental updates
 - d. 60 seconds
7. When a distance vector routing protocol process suspects that a route in its routing table is no longer valid, it will set the metric for the suspect route so high that the route is rendered unusable. This high metric is propagated in routing updates and sent out immediately. What is this process called?
- a. Count to infinity
 - b. Triggered updates
 - c. Split horizon
 - d. Feasible successor
8. What is the algorithm used by EIGRP?
- a. Bellman Ford
 - b. Dijkstra
 - c. DUAL
 - d. Administrative distance
9. If there is no path in the routing table, what action will EIGRP take?
- a. The process will poll the neighbors.
 - b. An ICMP redirect is sent to the router.
 - c. The route is dropped and a poison reverse is sent.
 - d. The route is given a metric of 255.
10. Which of the following has the best administrative distance?
- a. IGRP
 - b. RIPv1

- c. EIGRP
 - d. RIPv2
11. Which of the following criteria would select a route as the preferred path to be added to the routing table when there are multiple IP routing protocols running on the router?
- a. The route is the only available path.
 - b. The route has the one interface configured for IP.
 - c. The route has the lowest administrative distance.
 - d. The route has the lowest metric.
12. Which of the following statements are true about administrative distance?
- a. Administrative distance is the mechanism used by the routing process to select a path offered by two or more different routing protocols.
 - b. Administrative distance is a value given to each IP routing protocol.
 - c. Administrative distance can be manually configured.
 - d. The path that is chosen will be the one offered by the routing protocol with the lowest administrative distance.
13. When either RIPv1 or IGRP receive a flash update, what action do they take regarding the affected route?
- a. Place the affected route in holddown
 - b. Delete the affected route
 - c. Continue to use the route until they hear that it is unavailable
 - d. If a feasible successor is available, use it and delete the suspect route
14. In EIGRP, where does the routing process look for a feasible successor?
- a. Neighbor table
 - b. Routing table
 - c. Link state database
 - d. Topology table
15. If EIGRP finds a feasible successor, how long will it take for the network to converge?
- a. Within three routing updates
 - b. Almost instantly

- c. After the holddown timer has expired
 - d. After a routing update confirms the feasible successor is valid
16. Which of the following are examples of interior routing protocols?
- a. BGP-4
 - b. IS-IS
 - c. OSPF
 - d. RIPv1
17. Which of the following are examples of exterior routing protocols?
- a. BGP-4
 - b. IS-IS
 - c. OSPF
 - d. RIPv1
18. Autonomous system numbers are provided by which organization?
- a. ISO
 - b. IANA
 - c. OSI
 - d. IEEE

The answers to the “Do I Know This Already?” quiz are found in Appendix A, “Chapter ‘Do I Know This Already?’ Quizzes and Q&A Sections.” The suggested choices for your next step are as follows:

- **10 or less overall score** —Read the entire chapter. This includes the “Foundation Topics” and “Foundation Summary” sections, and the “Q&A” section at the end of the chapter.
- **11–15 overall score** —Begin with the “Foundation Summary” section and then go to the “Q&A” section at the end of the chapter. If you have trouble with these questions, read the appropriate sections in “Foundation Topics.”
- **16 or more overall score** —If you want more review on these topics, skip to the “Foundation Summary” section and then go to the “Q&A” section. Otherwise, move to the next chapter.

Foundation Topics

Distance Vector Routing Protocols

IP routing protocols use two main technologies: *link-state* and *distance vector* technologies. This chapter deals with distance vector protocols. You will learn about link-state protocols in Chapter 5, “IP Link-State Routing Principles.”

Distance vector protocols are the first routing protocols that were designed and include RIPv1 and IGRP. These protocols are classful protocols designed for small networks. As the networks started to expand, enhancements were made to the distance vector protocols, with the creation of RIPv2 and EIGRP. Although these still use the distance vector routing algorithms, they were developed later and are examples of classless routing protocols.

NOTE Although IGRP and EIGRP are defined by Cisco as distance vector routing protocols, they are proprietary protocols and therefore do not conform to all the characteristics of a distance vector protocol. For example, EIGRP is defined by Cisco as an advanced distance vector routing protocol, though it has adopted some of the link-state characteristics instead of the distance vector solution. It is sometimes referred to as a hybrid routing protocol. For the purposes of the exam, both IGRP and EIGRP are distance vector routing protocols.

Operation of Distance Vector Routing Protocols

Distance vector protocols send periodic updates about the networks that the routing process has discovered and entered into its routing table. The update is sent to directly connected neighbors. The destination address of the distance vector periodic update is 255.255.255.255 (the broadcast address), meaning that all systems on the segment hear the update. The update is sent out periodically after a timer has expired. The timer is reset immediately after the router has sent an update. Thus, the routing protocol sends out the entire routing table to its neighbors, sets a clock, and after a predetermined amount of time (30 seconds in the case of RIPv1), sends out the entire routing table again.

After receiving a neighbor’s routing table, the router updates its own table and sends the modified table in subsequent updates. Because the router passes on information that it has heard from its neighbors, as opposed to an update propagated from one router to all other routers in the network, distance vector routing protocols are said to be “routing by rumor.”

Distance vector routing protocols are classful routing protocols, meaning that they naturally summarize at the IANA or major network boundary. They do this by following the first octet rule.

The purpose of the distance vector protocol is to provide accurate, loop-free information to the routers. The update includes the entire routing table, excluding those networks that were learned from the interface through which the update is being sent. Thus, if a router knows of networks 1, 2, 3, 4, and 5, and it heard of networks 2 and 3 from an update received through E0, the routing update sent out of E0 would contain information about networks 1, 4, and 5. The rule of not repeating what you have heard back to the source is more than simply refined etiquette. It is the implementation of the *split horizon rule*. The purpose of this rule is to reduce network overhead and also prevent information from traveling in circles through the network, which can create *routing loops*.

To prevent routing loops, distance vector routing protocols employ the following techniques, which are described in more depth in the section “RIPv1”:

- Split horizon
- Poison reverse
- Holddown
- Triggered updates
- Aging of routes from the routing table

Count to infinity is a condition that occurs when a route vanishes from the network. It is found in distance vector routing protocols, although strictly speaking, it is not a technique.

The Distance Vector Routing Metrics

The metric used by distance vector protocols is often stated as being distance measured in the number of *hand-off points* or *hops* (routers) encountered on the way to the end device. Cisco defines IGRP and EIGRP as distance vector routing protocols. This muddies the original definition, because IGRP and EIGRP, although they are distance vector routing protocols, do not use hop count to determine the metric for a route; instead, they use several variables to compute the best path to a destination network. This is referred to as a composite and complex metric.

In distance vector routing protocols, the path selection is made using the Bellman Ford algorithm based on the metric or value of each available path, that of hop count. RFC 1058 discusses this in depth in reference to RIPv1. EIGRP, however, uses a proprietary algorithm called *Diffusing Update Algorithm (DUAL)*.

NOTE If you are asked a question on distance vector metrics, it might be wise to use the RFC 1058 definition of hop count because IGRP and EIGRP are proprietary protocols. Cisco also uses the original definition in its documentation.

RIP Version 1 and RIP Version 2

The next sections briefly describe the distance vector routing protocols RIPv1 and RIPv2.

RIPv1

Routing Information Protocol version 1 (RIPv1) is a simple routing protocol and, as such, works well in small networks that are unlikely to grow or change significantly. As a distance vector routing protocol, it sends updates every 30 seconds. These updates comprise the entire routing table.

Because RIPv1 was the first distance vector routing protocol and became an open standard, administrators who subsequently wrote a routing protocol were obliged to conform to the existing structure if they wished to define their routing protocol as a distance vector routing protocol. Therefore, the following characteristics describe not only RIPv1, but in essence, any distance vector routing protocol.

- **Count to infinity** — A router (A) hears about networks from its neighbors (B and C) and updates the routing table with the new networks. The routing table is then sent to all neighbors (B and C). However, if a neighbor (B) is sent information about networks for which it was the originating source, confusion can occur, which is referred to as a *routing loop*. The problem occurs when the path to a network goes down; each router might believe that there is an alternative path through its neighbor.

The ramifications of this problem are limited because each router increments the hop count before it sends out the update. When the hop count reaches 16, the network is rejected as unreachable, because the diameter of a RIPv1 network cannot be greater than 15. This is called *counting to infinity*, where “infinity” equals 16. Although the liability is controlled, it will still slow convergence of the network.

- **Split horizon** — This is a mechanism to prevent loops. If split horizon works, the need for “count to infinity” is eliminated. The split horizon rule states that the routing process will not advertise networks out of the interface through which those networks were heard. This prevents information about networks being repeated to the source of those networks.
- **Split horizon with poison reverse** — Split horizon on its own might not prevent loops, though it prevents networks being advertised out of the interface from which they were learned. However, poison reverse overrides split horizon when a network is lost. Poison reverse includes all the networks that have been learned from the neighbor, but it sets the metric to infinity (16). By changing the metric value to 16, the networks are reported to be unreachable. The routing process acknowledges the network but denies a valid path. Although this increases network overhead by increasing the update size, split horizon with poison reverse can prevent loops.

- **Holddown** — After deciding that a network in the routing table is no longer valid, the routing process waits for three routing updates (by default) before it believes a routing update with a less-favorable metric. Again, this is to prevent routing loops from generating false information throughout the network.
- **Triggered updates** — As soon as a routing process changes a metric for a network in its routing table, it sends an update with the metric set to a value that states it is unusable. In RIP, this value is infinity, that is, 16. Triggered updates inform the other routers immediately. If there is a problem in the network, all the affected routers go into holddown immediately instead of waiting for the periodic timer. This mechanism increases convergence and helps prevent loops.
- **Load balancing** — If the routing process sees multiple paths of equal cost to a remote network, it distributes the routed (datagram) traffic evenly among the paths. It will allocate datagrams to the different paths on a round-robin basis. The type of switching that is used—process switching or fast switching—will determine whether the load balancing is done on a round-robin or session basis. Round-robin load balancing is used when there is process switching in effect.

NOTE Because the metric used is hop count, one path can become saturated. A 56-kbps line and a 100-Mbps Fast Ethernet line might both offer paths of equal hop count; the user traffic would be divided equally between the links, but their bandwidths are disproportionate, allowing the slower link to become overwhelmed.

Cisco has implemented all the preceding options, which are defined in RFC 1058.

RIPv1 is useful in small networks and is distributed with Berkeley Standard Distribution (BSD), which makes it widely available. It might not be suitable for large environments, however, because the protocol was not designed with the expectation of being used in huge organizations. One of the keys to RIP is an understanding that 16 is the magic number.

As the network grows, you will see problems with applications timing out and congestion occurring on the network as the routers fail to adapt quickly to changes. When there has been a change in the network, the time that it takes for every router to register that change is known as the convergence time. The longer the convergence time, the greater the likelihood of problems on the network. Therefore, you need either to contain the growth of the network or to use a routing protocol that scales to a larger size. Open Shortest Path First (OSPF) is a link-state routing protocol that is designed to scale. It has the added advantage of being defined by the Internet Engineering Task Force (IETF), making it an industry standard in the public domain.

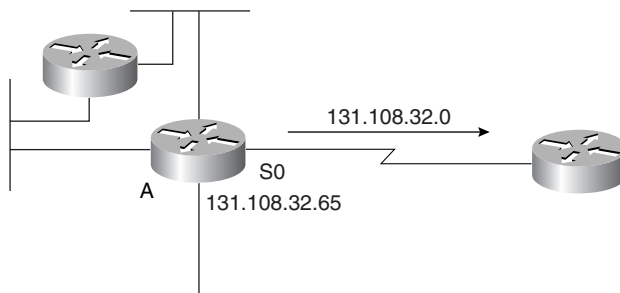
RIPv2

RIPv2 is an enhanced version of RIPv1. The main enhancements are that RIPv2 sends the subnet mask with the updates and that the updates are multicast, using 224.0.0.9. The following list explains in more detail the main characteristics of RIPv2.

- Using a multicast address saves network resources, because all the nodes in the network can discard the packet at either Layer 2 or 3 instead of taking the packet all the way up to the transport layer, where the port number is rejected. Although this saves system resources, the same amount of bandwidth is consumed on each link that the multicast traverses.
- Sending the mask in the updates allows the network to support classless interdomain routing (CIDR), summarization for the Internet, and variable-length subnet mask (VLSM) summarization for the autonomous system. Summarization requires manual configuration at either the routing process or the interface.

Figure 4-1 illustrates the configuration for Router A in Example 4-1.

Figure 4-1 RIPv2 Configuration



The following example shows a sample configuration, with the version of RIP defined at the routing process. You can also define the RIP version at the interface level if greater granularity is needed. The highlighted commands illustrate how RIP is configured as version 2 and how to manually summarize at the interface level.

Example 4-1 RIPv2 Configuration

```
RouterA#Config Term
RouterA(config)>enable
RouterA(config)#
RouterA(config)#router rip
RouterA(config-router)#version 2
RouterA(config-router)#network 131.108.0.0
RouterA(config-router)#no auto-summary
RouterA(config)#int s0
RouterA(config-if)#ip address 131.108.32.65 255.255.255.224
RouterA(config-if)#ip summary-address rip 131.108.32.0 255.255.224.0
```

IGRP and EIGRP

The next sections briefly describe the distance vector routing protocols IGRP and EIGRP.

IGRP

Internet Gateway Routing Protocol (IGRP) is a distance vector routing protocol created by Cisco Systems in the mid-1980s. It is a distance vector routing protocol, but because it is proprietary, it has the advantage of being capable of improving many of the elements seen in RIPv1, including incremental updates, fewer network resources to maintain the routing protocol, a more complex and efficient metric, and no limitation in diameter (total hop count) of the network.

As a proprietary protocol, IGRP can streamline its operation as a routing protocol, as described previously. A standard solution needs to implement all the options because by definition, it has to be all things to all people. The downside of a Cisco proprietary solution is that it can be implemented only on Cisco routers. However, IGRP is very efficient at sharing its information with other routing protocols using redistribution.

IGRP was designed to overcome the limitations of RIPv1. These enhancements center on the speed of convergence, a sophisticated routing metric, and ensuring that the best path is selected and entered into the routing table. The use of triggered updates in IGRP reduces the network resources that are required, which allows for larger networks and, thus, network growth. The following list summarizes the characteristics of IGRP and demonstrates that despite the fact that it is proprietary, IGRP is nonetheless a distance vector routing protocol:

- **Periodic updates** — In the spirit of a distance vector routing protocol, IGRP exchanges updates. IGRP sends updates every 90 seconds (by default), whereas RIP sends updates every 30 seconds. This update is a summary of the routing table and is exchanged only with adjacent routers.
- **Broadcasting updates** — Because the updates are sent to adjacent routers on the same segment, it is convenient to send these updates with a broadcast address. This means that all systems running IGRP on that segment will hear the update. Some protocols developed after IGRP, such as RIPv2, use multicast addressing to reduce the network overhead, which allows the network to scale.
- **Full routing table updates** — Although IGRP sends triggered updates whenever there is a change in the network, it also sends the full routing table every 90 seconds (by default) to ensure that all the routers have the same information. It is essential that the routers have the same information to avoid routing loops and to send packets via the best path. This allows other organizations or departments to merge or for a company to scale its network.

- **Count to infinity** —Like other distance vector routing protocols, IGRP uses the count to infinity to remove bad routes from the routing table. If the hop count to a known route increases, the route is removed from the routing table. The assumption is that another router has poisoned the route.
- **Split horizon** —Split horizon should prevent routing loops between adjacent routes. IGRP, like other distance vector routing protocols, believes that information about a route should never be sent back out of an interface through which it was heard. The theory is that if the updates are sent as broadcast, then every router on that segment has already received the information about that route. This not only prevents routing loops, but also reduces the size of the update by preventing redundant information from being included.
- **Triggered updates with route poisoning** —Distance vector routing protocols attempt to converge the network as quickly as possible. IGRP sends out an update as soon as a change in the network is noticed. If a route has failed, the hop count for that route is set to unreachable. Route poisoning is used in this way to break larger routing loops.
- **Load balancing on equal paths** (up to four paths, by default)— This is a Cisco feature that uses the network to achieve the greatest efficiency.
- **Default routes** —IGRP uses default routes in a slightly different manner from the traditional distance vector routing protocol. Although IGRP accepts the configuration of a default route, it also flags candidates for a default route. The candidate is identified as a route on the edge of the network. The candidate with the lowest metric is selected as the default route.
- **Bellman Ford routing algorithm** —This algorithm is the core of distance vector routing protocols, and IGRP modifies the algorithm in three important ways:
 - It uses a composite metric.
 - It is possible to configure more than one path to a destination. These are equal cost paths by default, but the **variance** command can be used to load balance across unequal paths.
 - The use of a composite metric uses more than one path to a destination, and triggered updates provide greater stability in the event of network change.

IGRP differs from RIPv1 in the following ways:

- The metric is a composite calculated from bandwidth, delay, loading, reliability, and maximum transmission unit (MTU). In fact, although MTU was originally designed as part of the metric, it is tracked but not used in the calculation. It is possible to configure the use of all the calculated elements of the metric. If all elements are not configured, the system will use only bandwidth and delay by default.
- The hop count is 100, configurable to 255 (although the hop count is not used as a metric, but to age out datagrams).
- The update timer is set by default to 90 seconds (three times that of RIPv1).

- Unequal-cost load balancing occurs on multiple paths.
- A more efficient packet structure is used.
- Autonomous systems are used to allow multiple processes within a routing domain, which allows the network to scale.

EIGRP

Enhanced Interior Gateway Routing Protocol (EIGRP) is a proprietary interior gateway protocol that is based on IGRP. EIGRP strives to reduce the need for network resources while decreasing the convergence times.

EIGRP is a proprietary protocol from Cisco and is sometimes referred to as an advanced distance vector protocol, an enhanced distance vector protocol, or a hybrid protocol. The term in current favor is advanced distance vector protocol.

EIGRP diverges from a classic distance vector routing protocol in many ways, solving many limitations in scaling the network. EIGRP has the following characteristics:

- In a stable network, EIGRP uses few resources; using only hello packets to maintain the routing databases.
- Changes in the network topology result in propagating only the changes instead of broadcasting the entire routing table every 30 seconds, whatever the status of the network.
- In EIGRP, the initial update is the complete routing table with subsequent updates carrying only changes when those changes occur. A neighbor table is built from hellos, ensuring that both the neighbor and the link are up and running. If the neighbor table fails to hear a hello in the allotted time, the routing process sends an update to the other neighbors informing them of a change in the network. The neighbor table also keeps track of acknowledgments (ACKs) received for the updates that have been sent. Reliable updates mean EIGRP does not have to retransmit every 30 seconds like RIPv1.
- Using Diffused Update Algorithm (DUAL), the network converges quickly with a minimum of network traffic.
- Instead of having to wait for updates, as soon as a route goes down, EIGRP examines the topology table for an alternative route. If an appropriate route is found, it is immediately placed into the routing table. In the event that no alternative path is found in the routing table, the neighbors are polled for alternative routes.

Choosing Between Routing Protocols Using Administrative Distance

Clearly, there are many IP routing protocols from which to choose. Choosing a single routing protocol is better, because the resulting consistency relates directly to the strength of the network. The network is complicated when more than a single routing protocol attempts to perform the same job.

When more than one routing protocol is running on the router, the routing process must make a decision to have one entry per prefix in the routing table. The choice cannot be based on the metric because metrics differ between routing protocols. Instead, another method, called *administrative distance*, was devised to solve the problem.

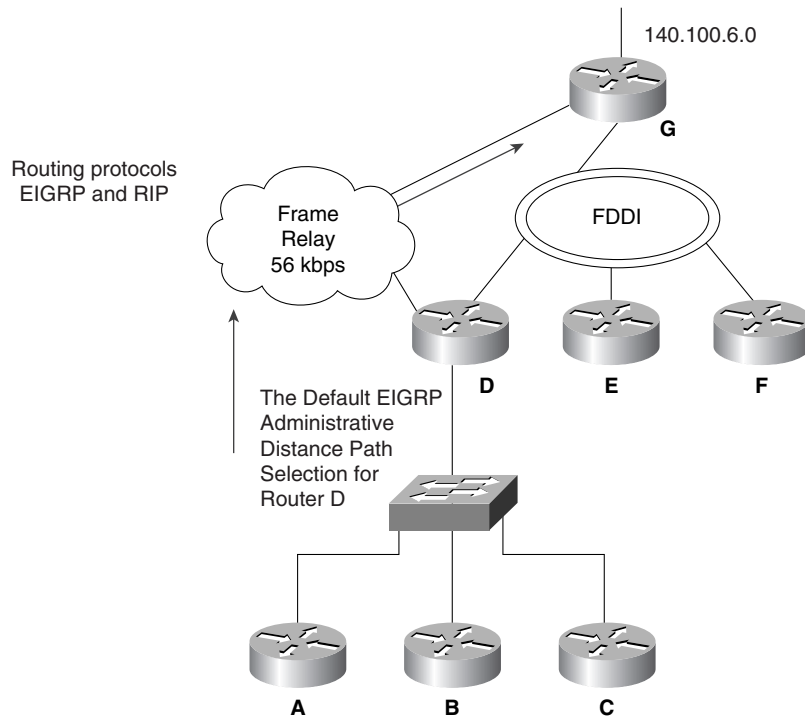
NOTE The routing table on a router running more than one routing protocol knows about all the networks heard by the various protocols and sends data to all of the distant networks, choosing the best path via administrative distance. However, a routing protocol only sends updates about networks it has knowledge of, so if IGRP knows about networks 1, 2, and 3, it propagates knowledge of these networks out of IGRP-configured interfaces to other IGRP routers on the same segment. It will not send out information about networks 4, 5, and 6 that were placed into the routing table by RIP.

In order for the IGRP routers to hear of networks 4, 5, and 6, it is necessary to share the network information between the routing protocols. This is called *redistribution*. However, the router that is responsible for redistribution will have more than one process running, which takes extra resources.

The administrative distance selects one or more paths to enter the routing table from several paths offered by multiple routing protocols.

In Figure 4-2, for example, both RIP and EIGRP have paths to the network 140.100.6.0. RIP is configured on the FDDI ring and EIGRP is running on the rest of the network. On Router D, RIP is offering a metric of 2 hops, and EIGRP is offering a metric of 768. Without redistribution, no conversion or choice is possible, because there are no similar criteria for distinguishing the two paths. Therefore, the metric is ignored, and the administrative distance is used to make the selection. The administrative distance of EIGRP is lower than that of RIPv1, so the path advertised by EIGRP is chosen, despite the speed of Frame Relay set at 56 kbps as opposed to the 100 Mbps of FDDI. In this case, if it is not possible to run EIGRP on the FDDI ring because of proprietary restrictions, manually configuring the administrative distance on Router D would be advisable.

Figure 4-2 *Path Selection Using Administrative Distance*



Administrative distance is a rather arbitrary set of values placed on the different sources of routing information. You can change the defaults, but proceed carefully when subverting the natural path selection. You must perform any manual configuration with careful reference to the network design of the organization and its traffic flow. The creation of floating static routes is an example of when the administrative distance is changed.

A lower administrative distance reflects the preferred choice. Table 4-2 lists the administrative distance defaults.

Table 4-2 *Default Administrative Distance*

Routing Source	Administrative Distance
Connected interface or static route that identifies the outgoing interface rather than the next hop	0
Static route	1
EIGRP summary route	5
External BGP	20

(continues)

Table 4-2 Default Administrative Distance (Continued)

Routing Source	Administrative Distance
EIGRP	90
IGRP	100
OSPF	110
RIP	120
External EIGRP	170
Internal BGP	200
An unknown network	255 or infinity

The administrative distance is looked at with total disregard of the metrics, which might result in a poor path selection. Problems can occur when redundancy is built into the network. For example, a low-cost, low-speed connection to a network can be used as a backup link to the core of the network or the segment that has the servers. The intention is for the link never to be used. The link is there as insurance against the primary link failing. Backup links for redundancy are often implemented over on-demand serial connections where the network charges are based on usage. However, in Figure 4-2, you have seen that this backup link would become the primary link.

To make this truly a backup link, you must configure it as a static route. However, the administrative distance of a static route takes precedence over everything but a directly connected network. Therefore, you need to configure manually the static route’s administrative distance to ensure that it takes precedence over any other route information only when the primary route fails. This design is called a *floating static route*.

Convergence

Convergence occurs when all the routers in the routing domain agree on the routes that are available. Convergence time is the time that it takes for every router’s routing table to synchronize after there has been a change in the network topology.

You need to ensure that the time taken is as short as possible, because while the routers disagree on the available networks, they cannot route data correctly or efficiently.

Each routing protocol has a different method of updating the routing table. This affects convergence time. The following sections introduce new concepts by explaining the different protocol convergence methods. The sections show the relative merits of each approach. The concepts are explained in depth in the chapters that concentrate on the specific protocols. The terms shown in *italics* are defined in the final glossary at the end of the book.

RIPv1 and RIPv2 Convergence

The steps for RIPv1 and RIPv2 convergence are as follows:

1. When the local router sees a connected route disappear, it sends a *flash update* and removes the route entry from its table. This is called a *triggered update* with *poison reverse*.
2. The receiving routers send flash updates and put the affected route in *holddown*.
3. The originating router queries its neighbor for alternative routes. If the neighbor has an alternative route, it is sent; otherwise, the *poisoned route* is sent.
4. The originating router installs the best alternative route that it hears because it has purged the original routes.
5. Routers that are in holddown ignore the alternative route.

When the other routers emerge from holddown, they will accept the alternative route.

Convergence takes the time for detection, plus holddown, plus the number of routing updates (equal to the hop-count diameter of the network).

IGRP Convergence

The steps for IGRP convergence are as follows:

1. When the local router sees a connected route disappear, it sends a flash update and removes the route entry from its table. This is called a triggered update with poison reverse.
2. The receiving routers send flash updates and put the affected route in holddown.
3. The originating router queries its neighbor for alternative routes. If the neighbor has an alternative route, it is sent; otherwise, the poisoned route is sent.
4. The originating router installs the best alternative route that it hears because it has purged the original routes. It sends a new flash update. This is the routing table, either with or without the network available, stating the higher metric.
5. Routers that are in holddown ignore the alternative route.
6. When the routers come out of holddown, they accept the alternative route.

When the other routers emerge from holddown, they will accept the alternative route.

Convergence takes the time for detection, plus holddown, plus the number of routing updates (equal to the hop-count diameter of the network). Because the time between updates is 90 seconds, this could take a very long time.

EIGRP Convergence

The steps for EIGRP convergence are as follows:

1. When the local router sees a connected route disappear, it checks the *topology table* for a feasible successor. If no *feasible successor* exists, it moves into active state.
2. The originating router queries its neighbor for alternative routes, and the receiving router acknowledges.
3. If an alternative route exists, information about this route is sent to the querying router.
4. If the router receives an acceptable successor, it adds the route to the table.
5. The router sends out a flash update of the path with the higher metric.
6. The receiving router acknowledges the update.

Convergence is quick because it is the detection time plus query time, reply time, and update time. If there is a feasible successor, convergence is almost instantaneous.

Interior and Exterior Gateway Protocols

Routing protocols that operate within an organization are referred to as Interior Gateway Protocols (IGPs) or *interior routing protocols* (for example, RIPv1, RIPv2, IGRP, EIGRP, OSPF, and IS-IS).

The boundaries of the organization are defined as the *autonomous system*. The unique number assigned to the autonomous system then identifies the organization. The autonomous system number might be viewed as another layer of hierarchy in the IP addressing scheme, because the number can represent a collection of IANA numbers.

Routing protocols that exchange routing information between organizations are known as Exterior Gateway Protocols (EGPs). EGPs are highly complex. The complexity arises from the need to determine policies between different organizations. Border Gateway Protocol Version 4 (BGP-4) is the only current example of an EGP.

Foundation Summary

The “Foundation Summary” section of each chapter lists the most important facts from the chapter. Although this section does not list every fact from the chapter that will be on your exam, a well-prepared candidate should, at a minimum, know all the details in each “Foundation Summary” before going to take the exam.

The following characteristics describe not only RIPv1, but in essence, any distance vector routing protocol.

- **Count to infinity** — A router (A) hears about networks from its neighbors (B and C) and updates the routing table with the new networks. The routing table is then sent to all neighbors (B and C). However, if a neighbor (B) is sent information about networks for which it was the originating source, confusion can occur, referred to as a routing loop. The problem occurs when the path to a network goes down; each router might believe that there is an alternative path through its neighbor.

The ramifications of this problem are limited because each router increments the hop count before it sends out the update. When the hop count reaches 16, the network is rejected as unreachable because the diameter of a RIPv1 network cannot be greater than 15. This is called counting to infinity, where “infinity” equals 16. Although the liability is controlled, it will still slow convergence of the network.

- **Split horizon** — This is a mechanism to prevent loops. If split horizon works, the need for “count to infinity” is eliminated. The split horizon rule states that the routing process will not advertise networks out of the interface through which those networks were heard. This prevents information about networks being repeated to the source of those networks.
- **Split horizon with poison reverse** — Split horizon on its own might not prevent loops, though it prevents networks being advertised out of the interface from which they were learned. However, poison reverse overrides split horizon when a network is lost. Poison reverse includes all the networks that have been learned from the neighbor, but it sets the metric to infinity (16). By changing the metric value to 16, the networks are reported to be unreachable. The routing process acknowledges the network but denies a valid path. Although this increases network overhead by increasing the update size, split horizon with poison reverse can prevent loops.
- **Holddown** — After deciding that a network in the routing table is no longer valid, the routing process waits for three routing updates (by default) before it believes a routing update with a less-favorable metric. Again, this is to prevent routing loops from generating false information throughout the network.

- **Triggered updates** — As soon as a routing process changes a metric for a network in its routing table, it sends an update with the metric set to a value that states it is unusable. In RIP, this value is infinity, that is, 16. Triggered updates inform the other routers immediately. If there is a problem in the network, all the affected routers go into holddown immediately instead of waiting for the periodic timer. This mechanism increases convergence and helps prevent loops.
- **Load balancing** — If the routing process sees multiple paths of equal cost to a remote network, it distributes the routed (datagram) traffic evenly among the paths. It will allocate datagrams to the different paths on a round-robin basis. The type of switching that is used—process switching or fast switching—will determine whether the load balancing is done on a round-robin or session basis. Round-robin load balancing is used when there is process switching in effect.

Table 4-3 summarizes default administrative distances.

Table 4-3 *Default Administrative Distance*

Routing Source	Administrative Distance
Connected interface or static route that identifies the outgoing interface rather than the next hop	0
Static route	1
EIGRP summary route	5
External BGP	20
EIGRP	90
IGRP	100
OSPF	110
RIP	120
External EIGRP	170
Internal BGP	200
An unknown network	255 or infinity

Q&A

As mentioned in the introduction, “All About the CCNP, CCDP, and CCIP Certifications,” you have two choices for review questions. The questions that follow give you a bigger challenge than the exam itself by using an open-ended question format. By reviewing now with this more difficult question format, you can exercise your memory better and prove your conceptual and factual knowledge of this chapter. The answers to these questions are found in Appendix A.

For more practice with examlike question formats, including questions using a router simulator and multichoice questions, use the exam engine on the CD-ROM.

1. Name one routing protocol that sends periodic updates.
2. What is an incremental update, and how often is it sent out?
3. Distance vector routing protocols naturally summarize at which boundary?
4. What is the algorithm used by distance vector protocols?
5. Give three reasons why RIPv1 has problems working in a large network.
6. What is the destination address of the distance vector periodic update in RIPv1?
7. State two ways that a route is selected as the preferred path.
8. What is administrative distance?
9. If IGRP has three paths to a remote network in which each path has an equal metric, what will happen?
10. A distance vector routing protocol uses the mechanism of poison reverse. What is poison reverse?
11. Name two distance vector routing protocols.
12. Describe the mechanism of split horizon.
13. What is meant by the phrase *routing by rumor*?
14. Why does the use of multicast addressing in RIPv2 overcome some of the limitations of RIPv1?
15. Explain the use of holddown in distance vector routing protocols to create stability in the network.
16. What is the maximum hop count in RIPv1 and RIPv2?
17. Both EIGRP and IGRP use a composite metric. What are the main components of this metric?
18. Explain briefly how RIPv2 differs from RIPv1.
19. What is meant by the term *convergence*?
20. Give the configuration commands to turn on the process for RIPv2.



This chapter covers the following topics, which you need to understand to pass the CCNP/CCDP/CCIP BSCI exam:

- Link-state routing protocol overview
- OSPF
- IS-IS
- BGP-4
- Convergence

IP Link-State Routing Principles

In this chapter, the concepts of routing with IP using the link-state algorithm and the mechanics of the process are dealt with generically as a foundation for the subsequent chapters. Although this chapter compares the routing protocols OSPF, IS-IS, and BGP-4, subsequent chapters deal with each routing protocol individually. The subsequent chapters assume that you comprehend the subjects covered in this chapter.

“Do I Know This Already?” Quiz

The purpose of the “Do I Know This Already?” quiz is to help you decide what parts of this chapter to use. If you already intend to read the entire chapter, you do not necessarily need to answer these questions now.

The 15-question quiz, derived from the major sections in the “Foundation Topics” portion of the chapter, helps you to determine how to spend your limited study time.

Table 5-1 outlines the major topics discussed in this chapter and the “Do I Know This Already?” quiz questions that correspond to those topics.

Table 5-1 “Do I Know This Already?” Foundation Topics Section-to-Question Mapping

Foundation Topics Section	Questions Covered in This Section
Link-State Routing Protocol Overview	1–3
OSPF	4–6
IS-IS	7–9
BGP-4	10–12
Convergence	13–15

NOTE The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark this question wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

1. In a network running a link-state routing protocol, every router will have which of the following?
 - a. A localized routing table
 - b. An identical view of the network
 - c. A spanning tree
 - d. TCP connections with adjacent neighbors
2. Link-state routing protocols use which of the following algorithms?
 - a. Bellman Ford
 - b. Dijkstra
 - c. DUAL
 - d. Path attributes
3. A link-state routing protocol uses which of the following?
 - a. Incremental updates
 - b. Hello packets
 - c. Topology databases
 - d. Transport layer protocols
4. Which of the following are supported by OSPF?
 - a. VLSM
 - b. Split Horizon
 - c. Path Attributes
 - d. Classless routing
5. OSPF uses a hierarchical design for which of the following reasons?
 - a. To support VLSM
 - b. To prevent the 15-hop limitation

- c. To conserve network resources
 - d. To limit the scope of poison reverse
- 6. OSPF uses multicast addresses for updates. What are the multicast addresses?
 - a. 224.0.0.2
 - b. 224.0.0.4
 - c. 224.0.0.6
 - d. 224.0.0.5
- 7. Which Layer 3 protocols are supported by IS-IS?
 - a. IP, AppleTalk, and IPX
 - b. IP and CLNS
 - c. IP and IPV6
 - d. DECnet Phase V
- 8. IS-IS runs at which layer?
 - a. Layer 3
 - b. Layer 4
 - c. Layer 2
 - d. Layer 5
- 9. IS-IS is defined in which document?
 - a. IETF 1195
 - b. ISO 1195
 - c. RFC 2043
 - d. ISO 10589
- 10. Which of the following best describes the routing protocol BGP-4?
 - a. BGP-4 connects autonomous systems.
 - b. BGP-4 sends the entire routing table every 20 minutes.
 - c. BGP-4 uses path vectors to determine the best path.
 - d. BGP-4 is an interior routing protocol.

11. At what layer does BGP-4 send keepalives?
 - a. Layer 2
 - b. Layer 3
 - c. Layer 5
 - d. Layer 4
12. Which of the following characteristics is true of BGP?
 - a. Uses a best effort delivery
 - b. Sends the entire routing table every 30 minutes
 - c. Uses a holddown of 30 seconds
 - d. Sends only trigger or incremental updates after the initial setup
13. When an OSPF router receives an LSA, which of the following best describes the action taken?
 - a. Update the routing table and flood the new routing table out of all interfaces
 - b. Update the topology table and flood the LSA out of its interfaces
 - c. Put the suspect route into holddown for 30 seconds
 - d. Mark the entry as suspect and query neighbors for a new route
14. Which of the following routing protocols runs the Dijkstra algorithm to rebuild the routing table?
 - a. IS-IS
 - b. EIGRP
 - c. BGP
 - d. OSPF
15. When a neighbor is no longer available, what action will the BGP router take?
 - a. Clears the route from the routing table and floods a LSA out of all interfaces
 - b. Queries neighbors for a new route
 - c. Uses a holddown of 30 seconds
 - d. Tries to reconnect to its neighbor

The answers to this quiz are found in Appendix A, “Answers to Chapter ‘Do I Know This Already?’ Quizzes and Q&A Sections.” The suggested choices for your next step are as follows:

- **7 or less overall score** —Read the entire chapter. This includes the “Foundation Topics” and “Foundation Summary” sections, the “Q&A” section, and the “Scenarios” at the end of the chapter.
- **8–10 overall score** —Begin with the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. If you have trouble with these exercises, read the appropriate sections in “Foundation Topics.”
- **11 or more overall score** —If you want more review on these topics, skip to the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. Otherwise, move to the next chapter.

Foundation Topics

Link-State Routing Protocol Overview

A *link-state routing protocol* is a sophisticated protocol dedicated to maintaining loop-free, accurate routing tables. It does not send the entire routing table periodically via broadcasts, as the original distance vector protocols (such as RIPv1) do, but instead uses multicast addressing and incremental updates. Some routing protocols send incremental updates in addition to a compressed copy of the routing table. However, the full routing update is sent every 30 minutes, instead of every 30 seconds, and has a multicast address.

The Meaning of Link State

A *link* refers to the connection between routers, that is, the physical connection or medium between the routers, over which a logical link is formed. A link-state routing protocol is therefore a protocol that sends information about the links between routers, when there is a change in the state of one of those links. Thus, when the Ethernet connection between Router A and Router B fails, an update is propagated by Router A and Router B, informing the entire network that the link between A and B is in the down state.

Unlike distance vector protocols, the information concerns only the local links (not the routes) connected to the router, and these links are propagated, unchanged, to every other router in the network. Therefore, every router has the same image of the network, created from the original updates from every other router in the network. Sending an update about links is more efficient than sending data about routes, because one link might affect many routes. Sending information about the links allows the routers to compute the routes that might be affected. The resources used are router CPU rather than network bandwidth.

Learning About the Network

The routing protocol develops and maintains the neighbor relationship with routers on the same link by sending a simple hello message across the medium. This is a connection-oriented exchange. After the routers have synchronized their routing tables by exchanging routing updates, they are deemed to be adjacent neighbors.

This neighbor relationship and the subsequent adjacency is maintained as long as the Hello protocol is received. For this to work, the two routers must have the same subnet mask and hello timers. Because the neighbor relationship is continuous, information can be exchanged between the routing processes quickly and efficiently. Therefore, link changes in the network are realized very quickly.

A router knows quickly whether the neighbor, who might also be the next hop, is dead, because the router no longer receives Hello protocol messages.

As soon as the routing protocol identifies a problem, it sends out a message immediately, without waiting for the update timer to expire. This is also known as a *triggered update*. This is an *incremental update* because it contains only the network change. The incremental update improves convergence time and also reduces the amount of information that needs to be sent across the network. The network overhead on the physical media is eased, allowing more bandwidth for data.

Link-state routing protocols are used in larger networks because the method that they use to update the routing tables requires fewer network resources.

Link-state routing protocols attempt to reduce network overhead by:

- Using multicast addressing
- Sending triggered updates
- Sending network summaries infrequently, if at all
- Using small packets from every router to describe their local connectivity, instead of the entire routing table

Updating Local Network Tables

A link-state protocol holds a topology database, a network map of every link seen by the routing protocol. The topology database of the network updates the routing table database, after the incremental updates are received and processed. In OSPF, for example, the incremental updates are called link-state advertisements (LSAs). After an update is received and forwarded, the routing protocol computes a new topology database and, from this, a new path. The routing protocol uses the Dijkstra algorithm to achieve this new understanding of the network.

Path Selection

The routing protocol selects the best path to a destination, via the metric. Link-state routing protocols state the metric to be cost, although many vendors supply a default that can be overridden manually. This is true of Cisco's implementation of OSPF, which uses the inverse of bandwidth as its default.

Examples of link-state routing protocols for IP are OSPF and IS-IS.

OSPF

Rarely is a name as descriptive as the one given to this protocol, Open Shortest Path First (OSPF). OSPF is an open standard, defined in detail in many RFCs, including RFC 2328. OSPF uses the SPF algorithm to compute the best path to any known destination. OSPF ensures a loop-free topology with fast convergence, although it can use a lot of CPU. OSPF was devised to overcome the limitations of early distance vector protocols, such as RIPv1.

OSPF, as a link-state routing protocol, is an improvement over a distance vector routing protocol, such as RIPv1, for large networks for the following reasons:

- It uses bandwidth more efficiently by sending incremental updates while requiring greater memory and CPU to calculate the Dijkstra algorithm.
- The updates are not broadcast as in RIPv1 but are directed to multicast addresses 224.0.0.5 and 224.0.0.6.
- It propagates changes in the network more quickly with incremental updates and neighbor relationships.
- It is not limited in size by a maximum hop count of 15.
- It allows for variation in network design throughout the organization, using VLSM.
- It has security options, allowing it to use the Message-Digest version 5 (MD5) specification.
- The metric can be defined manually, allowing for greater sophistication in the path determination.
- It is more responsive to network changes and is flexible in network addressing and design, allowing the network to scale.

OSPF is designed to offer the greatest flexibility in network design. As an open standard, it is required to offer interoperability while allowing the network to grow. These requirements make OSPF a highly complex routing protocol.

To understand this complexity, it is useful to identify the main characteristics of OSPF. These key attributes of OSPF include the following:

- Maintains adjacent neighbors.
- Uses hello timers to maintain adjacencies. These are sent every 60 seconds on a WAN and every 10 seconds on LAN. If nothing has been heard from a neighbor within four times the hello timer, the neighbor is declared dead, requiring the generation of an LSA.
- Sends the minimum amount of information in an incremental update when there has been a change in the network. This allows for fast network convergence. If the network is stable and there have been no updates within 30 minutes, a compressed update is sent.
- Adds another level of hierarchy to the IP address by designing networks into areas.
- Is a classless routing protocol.
- Uses VLSM and both manual and automatic summarization at the IANA class boundary.
- Uses cost as the metric, defined by Cisco to be the inverse bandwidth; the formula is $10^8/\text{bandwidth}$ (in bps).
- Assigns specific functionality to different routers to streamline the process of communication change in the network.
- Operates within an organization as an interior routing protocol.

IS-IS

IS-IS and OSPF share many of the same features because they both attempt to solve the limitations in distance vector routing protocols. Like OSPF, IS-IS is a link-state routing protocol that uses the SPF routing algorithm. Both OSPF and IS-IS offer fast convergence, are flexible, and are designed to resist routing loops and to support very large networks.

IS-IS is an integrated protocol. First designed by Digital Equipment Corporation for DECnet Phase V, it became a standard ratified by the International Standards Organization (ISO). It has a large address space, allowing for incredibly large networks, such as those in the United States government, including the armed forces. The hierarchical design of the protocol allows for this large size in both the interpretation of the address and the transmission of the routing updates. The packet structure was conceived with the intention of allowing the protocol to incorporate enhancements, making it a very flexible protocol.

IS-IS has the following features:

- It routes CLNP traffic, as defined in the ISO 10589 standard.
- It routes IP traffic, as defined in RFC 1195.
- It is a classless routing protocol.
- It allows VLSM and both manual and automatic summarization at the IANA class boundary.
- It uses the network design of areas to limit CPU-intensive computation.
- It uses metric of cost defined by Cisco to be 10 on all media.
- It assigns functionality to routers to streamline the communication of network change. Level 1 routers deal with interarea updates, whereas Level 2 routers communicate between areas.
- It sends incremental updates to conserve both bandwidth and CPU, though broadcast media synchronize databases every 10 minutes.
- It maintains neighbor relations through the Hello protocol, sent every 10 seconds on all media.
- It considers neighbors dead after 30 seconds of silence.
- It operates within an autonomous system as an internal routing protocol.

BGP-4

BGP is not a link-state routing protocol. Strictly speaking, it is a path vector routing protocol, which has some of the characteristics of both link-state and distance vector routing protocols. It is an exterior routing protocol and, as such, is completely different from anything seen before. It is included in this comparison chapter on link-state routing protocols because it fits most conveniently here as one of the more complex protocols. The term *path vector* refers to the list of autonomous system numbers that are carried in the BGP-4 updates. The vector indicates the direction to send the

traffic to find the path to a remote network. Developed to connect an enormous amount of networks together, BGP is used primarily to connect the Internet and Internet service providers (ISPs).

There are two flavors of BGP: internal BGP (IBGP) and external BGP (EBGP). Essentially, BGP is an external routing protocol used to connect BGP autonomous systems, referred to as EBGP. IBGP is used to send routing information internally across an autonomous system, using it as a transit area to another autonomous system. IBGP needs a fully meshed BGP network, but the routers do not need to be directly connected. BGP updates can be sent to the other BGP routers, or the BGP data traffic can find the remote destination by listening to the interior IP routing protocol. Although the remote peer does not have to be directly connected, an entry must be in the routing table of the remote peer for the routers to communicate with each other.

BGP, which is defined in RFC 1771, sends very little information in its updates, which are only sent when there is a change in the network. One of the main goals of BGP is to allow you to determine the path that different types of traffic can take. It is essentially possible to program the routing protocol to allow traffic from one source to take the high road, while other traffic is sent on the low road. This flexibility and the ability to grow the network to large sizes are the main strengths of BGP. This is a very different protocol from the other protocols studied so far, as shown in the following list of characteristics:

- It is a classless routing protocol.
- It allows VLSM and both manual and automatic summarization.
- It sends full routing updates at the beginning of the session.
- It sends only trigger or incremental updates after the initial setup.
- It maintains connections between BGP routers by using periodic hellos every 60 seconds. After 180 seconds, the neighbor is declared dead. The Hello protocol is connection-orientated, using TCP, port 179.
- It uses the hierarchical structure of autonomous systems.
- It has a complex metric called *attributes* by which traffic paths can be manipulated.

Convergence

Each routing protocol has a different method of updating the routing table, affecting the time to converge the routing tables. Some new concepts are introduced in the following comparison. The concepts are explained in depth in the chapters that concentrate on the specific protocols.

OSPF Convergence

The steps for OSPF convergence are as follows:

1. When a router detects a link failure, the router sends an LSA to its neighbors. If the router is on a multiaccess link, it sends the update to the designated router (DR) and the backup designated router (BDR), not to all neighbors.
2. The path is removed from the originating router's tables.
3. On receipt of the LSA, all routers update the topology table and flood the LSA out its interfaces.
4. The routing protocol runs the Dijkstra algorithm to rebuild the routing table.

For OSPF, convergence is detection time, plus LSA flooding, plus 5 seconds before computing the topology table. This amounts to a few seconds.

IS-IS Convergence

The steps for IS-IS convergence are as follows:

1. When a router detects a link failure, an LSP is sent to its neighbors. If the router is on a multiaccess link, the update is sent to the designated intermediate system (DIS the IS-IS term for a designated router), not to all neighbors.
2. The path is removed from the originating router's tables.
3. On receipt of the LSP, all routers update the topology table and flood the LSP out its interfaces, except for the interface that received the LSP.
4. Each router runs the Dijkstra algorithm to rebuild the forwarding table.

For IS-IS, convergence is detection time, plus LSP flooding. The time to converge the network amounts to a few seconds. If convergence is deemed to be the topology table being updated, this could take longer.

BGP Convergence

BGP convergence is different, depending on whether IBGP or EBGP is being run. Reliability is far more important to EBGP than how long it takes to update the routing table, whereas IBGP needs to ensure a faster convergence to remain synchronized with the interior routing protocol.

When a neighbor is no longer available, the BGP router tries to reconnect to its neighbor. If this fails, the session is formally closed and the information from the router is removed from the database. An update is sent to all neighbors.

Foundation Summary

The “Foundation Summary” section of each chapter lists the most important facts from the chapter. Although this section does not list every fact from the chapter that will be on your exam, a well-prepared candidate should, at a minimum, know all the details in each “Foundation Summary” before going to take the exam.

Table 5-2 is a summary of IP routing protocols and the update timers.

Table 5-2 *A Summary of IP Routing Protocols and the Update Timers*

Protocol	Update Timer	Technology
RIPv1	Every 30 seconds for entire routing table.	Distance vector.
RIPv2	Every 30 seconds for entire routing table.	Distance vector.
OSPF	Incremental with only the network change. However, 30 minutes after the last update was received, a compressed version of the table is propagated.	Link state.
EIGRP	Incremental updates with network change only.	Advanced distance vector, sometimes called enhanced distance vector or a hybrid routing protocol.
IGRP	Updates every 90 seconds with incremental updates as needed.	Distance vector.
BGP-4	Incremental with only the network change.	Path vector, sometimes referred to as a type of distance vector routing protocol.
IS-IS	Incremental with only the network change. However, the router that originated the LSP must periodically refresh its LSPs to prevent the remaining lifetime on the receiving router from reaching 0. The refresh interval is 15 minutes. This means that approximately 15 minutes after the last update was received, a compressed list of all the links the router has knowledge of is sent to all routers.	Link state.

Table 5-3 summarizes the major differences between distance vector routing protocols and link-state routing protocols.

Table 5-3 *Distance Vector Routing Protocols Versus Link-State Routing Protocols*

Distance Vector	Link-State
Sends its entire routing table at periodic intervals out of all interfaces (typically, this is based in seconds). Sends triggered updates to reflect changes in the network.	Sends incremental updates when a change is detected. OSPF will send summary information every 30 minutes, regardless of whether incremental updates have been sent in that time.
Typically involves updates sent using a broadcast address to everyone on the link.	Typically involves updates sent to those routers participating in the routing protocol domain, via a multicast address.
Uses a metric based on how distant the remote network is to the router. (IGRP does not conform to this as a proprietary solution.)	Is capable of using a complex metric, referred to by OSPF and IS-IS as cost.
Has knowledge of the network based on information learned from its neighbors.	Has knowledge of the network based on information learned from every router in the area.
Includes a routing table that is a database viewed from the perspective of each router.	Has a topological database that is the same for every router in the area. The routing table that is built from this database is unique to each router.
Uses the Bellman Ford algorithm for calculating the best path.	Uses the Dijkstra algorithm.
Does not consume many router resources, but is heavy in the use of network resources.	Uses many router resources, but is relatively low in its demand for network resources.
Maintains one domain in which all the routes are known.	Has a hierarchical design of areas that allow for summarization and growth.
Is not restricted by addressing scheme.	For effective use, the addressing scheme should reflect the hierarchical design of the network.
Involves slower convergence because information of changes must come from the entire network (but indirectly). Each routing table on every intervening router must be updated before the changes reach the remote end of the network.	Involves quicker convergence because the update is flooded immediately throughout the network.

Table 5-4 summarizes the differences between RIPv1 and OSPF. RIPv1, as the first distance vector routing protocol, and OSPF, as the first link-state routing protocol, are very familiar to most in the networking industry and thus easily used as examples for comparison.

Table 5-4 *RIPv1 Versus OSPF*

RIPv1	OSPF
Is a simple protocol to design, configure, and maintain.	Is a complex protocol to design and, in some instances, to configure and maintain.
Does not require a hierarchical addressing scheme.	If full benefits of the protocol are to be harnessed, should use a hierarchical IP addressing scheme.
Does not pass the subnet mask in the routing update and therefore is not capable of classless routing or VLSM.	Carries the mask in the update and therefore can implement VLSM, summarization, and classless routing.
Is limited to a 15-hop diameter network.	Is unlimited in the diameter of the network, although it is suggested that an area not exceed more than 50 networks.
Does not acknowledge routing updates; just repeats them periodically (every 30 seconds).	Acknowledges updates.
Has a routing table that is sent out of every interface every 30 seconds (by default).	Involves updates sent as required (when changes are seen) and every 30 minutes after no change has been seen.
Can transmit information about the network in two messages: the routing update and the triggered update.	Has protocols for discovering neighbors and forming adjacencies, in addition to protocols for sending updates through the network. These protocols alone add up to nine message types.
Uses hop count as a metric, the number of routers to process the data.	Uses cost as a metric. Cost is not stated in the RFCs, but it has the capacity to be a complex calculation, as seen in Cisco's implementation.

Table 5-5 summarizes the major differences between all available IP routing protocols.

Table 5-5 *Comparison Chart for IP Routing Protocols*

	RIPv1	RIPv2	IGRP	EIGRP	OSPF	IS-IS	BGP
Distance Vector	✓	✓	✓	✓	–	–	Path vector
Link State	–	–	–	–	✓	✓	–
Classless	–	✓	–	✓	✓	✓	✓

Table 5-5 *Comparison Chart for IP Routing Protocols (Continued)*

	RIPv1	RIPv2	IGRP	EIGRP	OSPF	IS-IS	BGP
Classful	✓	–	✓	–	–	–	–
VLSM	–	✓	–	✓	✓	✓	✓
Manual Summarization	–	✓	–	✓	✓	✓	✓
Automatic Summarization at IANA boundary	✓	✓	✓	✓	–	–	✓
Metric	Hop	Hop	Composite	Composite	Cost	Cost	Path attributes
Max Hop Count	15	15	255	255	–	–	–
Update Timers	30 sec	30 sec + triggered	90 sec + triggered	Triggered	30 min + triggered	Synchronized every 15 min on broadcast media+ triggered	Triggered
Hello	None	None	None	60 sec <T1, everything else 5 sec	30 sec WAN, everything else 10 sec	10 sec	60 sec
Dead Time	180 sec, flush after 240 sec	180 sec, flush after 240 sec	3 * hello	3 * hello	4 * hello	Hold for 30 sec	180 sec

Q&A

As mentioned in the introduction, “All About the Cisco CCNP, CCDP, and CCIP Certifications,” you have two choices for review questions. The questions that follow next give you a bigger challenge than the exam itself by using an open-ended question format. By reviewing now with this more difficult question format, you can exercise your memory better and prove your conceptual and factual knowledge of this chapter. The answers to these questions are found in Appendix A.

For more practice with examlike question formats, including questions using a router simulator and multichoice questions, use the exam engine on the CD.

1. What is the routing algorithm used in OSPF?
2. State one method by which a link-state routing protocol attempts to reduce the network overhead.
3. What is the purpose of the Dijkstra algorithm?
4. Name two link-state IP routing protocols.
5. Name the TCP port used by BGP-4.
6. State the metric used by OSPF.
7. How often does Integrated IS-IS send out new LSAs?
8. State one way that OSPF is an improvement over RIPv1.
9. State one key attribute of OSPF.
10. State one key attribute of IS-IS.
11. State one key attribute of BGP-4.
12. What is the default hello update timer for IS-IS on broadcast media?
13. On a broadcast link, how long does OSPF wait by default before it determines that a neighbor is dead?
14. What is iBGP?
15. When does OSPF send updates?
16. When does BGP send updates?
17. What is a topological database?
18. What is an adjacent neighbor?
19. What is a triggered update?
20. What is required for iBGP to operate?

Scenarios

The following scenarios and questions are designed to draw together the content of the chapter and to exercise your understanding of the concepts. There is not necessarily a right answer. The thought process and practice in manipulating the concepts is the goal of this section. The answers to the scenario questions are found at the end of this chapter.

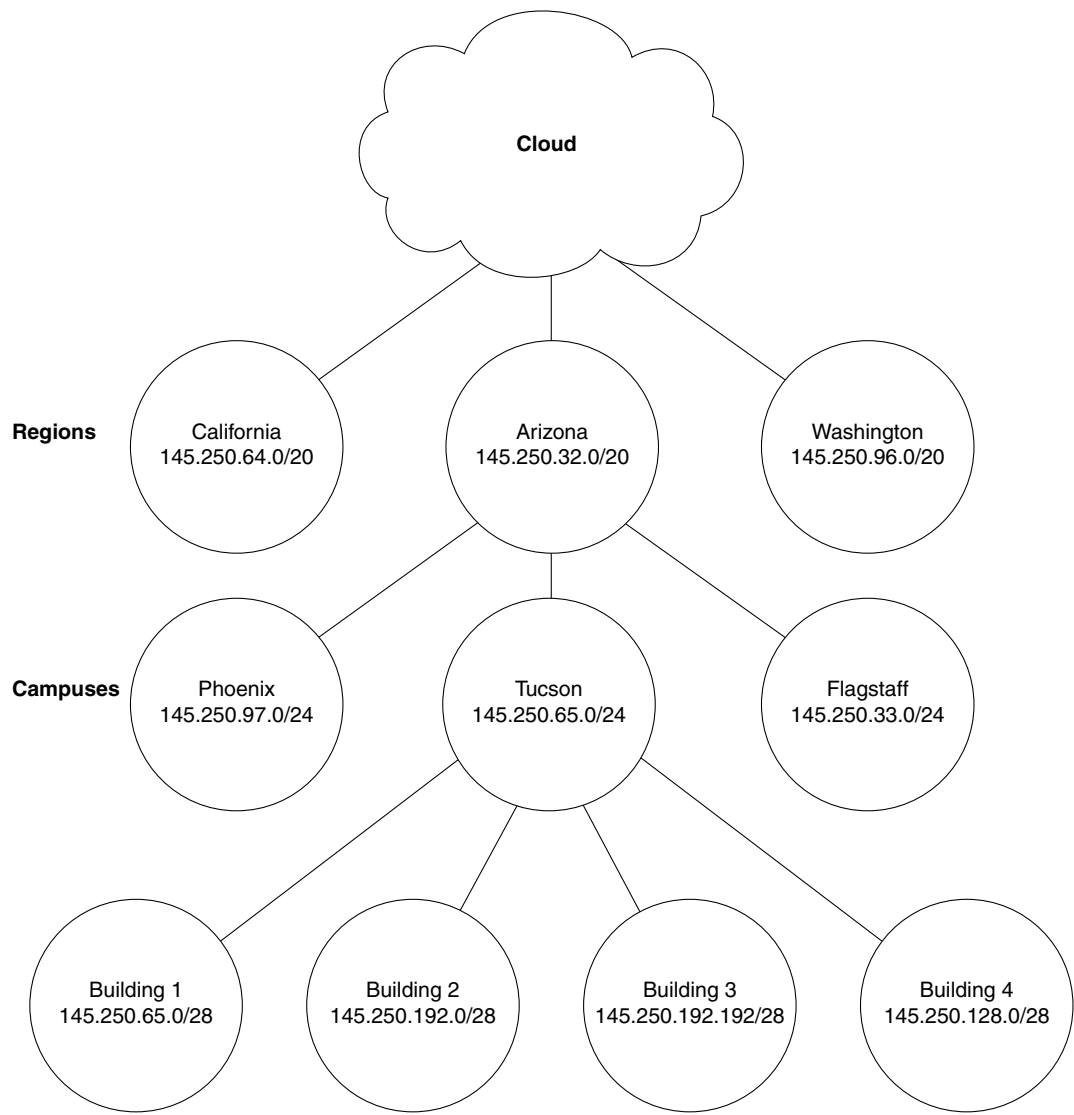
Scenario 5-1

Mental Merge, a company with many ideas but no bandwidth with which to develop or communicate them, is in need of a routed solution. Using the addressing scheme provided in Scenario Solution 3-1 in Chapter 3, “Designing IP Networks,” it is now necessary to implement routing.

Using the network in Figure 5-1, answer the following questions.

1. Using appropriate addressing and in reference to Figure 5-1, state where the routers should be placed.
2. The administrator has decided that a link-state routing protocol is the best solution for this network design. Justify this choice, explaining which characteristics of the link-state routing protocol would benefit this network.
3. The administrator must create an implementation plan for the team. List the IP routing protocol requirements for every router that might be used as a checklist for the installation staff.
4. The links between the various sites are leased lines with a backup link using a dialup line. Should the administrator be aware of any considerations?

Figure 5-1 The Mental Merge Company Network for Scenario 5-1



Scenario Answers

The answers provided in this section are not necessarily the only possible answers to the questions. The questions are designed to test your knowledge and to give practical exercise in certain key areas. This section is intended to test and exercise skills and concepts detailed in the body of this chapter.

If your answer is different, ask yourself whether it follows the tenets explained in the answers provided. Your answer is correct not if it matches the solution provided in the book but rather if it has included the principles of design laid out in the chapter.

In this way, the testing provided in these scenarios is deeper: It examines not only your knowledge but also your understanding and capability to apply that knowledge to problems.

If you do not get the correct answer, refer back to the text and review the subject tested. Be certain to also review your notes on the question to ensure that you understand the principles of the subject.

Scenario 5-1 Answers

1. *Using appropriate addressing and in reference to Figure 5-1, state where the routers should be placed.*

The routers should be placed in each location, with the option of adding routers within each building if the network grows considerably.

2. *The administrator has decided that a link-state routing protocol is the best solution for this network design. Justify this choice, explaining which characteristics of the link-state routing protocol would benefit this network.*

A link-state routing protocol would be a good choice because of the large number of WAN interfaces. A distance vector routing protocol would increase congestion across these low-bandwidth links. The capability to use VLSM and to summarize these points would be an added advantage.

3. *The administrator must create an implementation plan for the team. List the IP routing protocol requirements for every router that might be used as a checklist for the installation staff.*

Each person implementing the routing protocol on the router would have to ensure the following:

- The appropriate interfaces have IP addresses that are on the same subnet as the other devices on the segment.
- The routing protocol is configured correctly with the correct network addresses.

- The routing table reflects the logical topology map of the network and all the remote networks are present.
 - If there are multiple paths available of equal cost, the routing protocol should be load sharing between the paths. This means all the paths are present in the routing table.
4. *The links between the various sites are leased lines with a backup link using a dialup line. Should the administrator be aware of any considerations?*

The leased lines to the remote sites could be configured to be the primary link; as such, no traffic would traverse the dialup links. However, routing updates would be propagated out of the dialup links so that the routing table would be aware of the potential path. To prevent this (and, thus, the dialup line being raised), the path could be manually entered into the routing table. However, this would render it the preferred path. Configuring the dialup paths as floating static routes would ensure that they were used only if the primary line failed, without having to generate network traffic across the link to maintain the routing table.

Table 5-6 summarizes the major differences between distance vector routing protocols and link-state routing protocols.

Table 5-6 *Distance Vector Routing Protocols Versus Link-State Routing Protocols*

Distance Vector	Link-State
Sends its entire routing table (typically every 30 seconds).	Sends incremental updates. It synchronizes the routing tables every 15 or 30 minutes.
Updates sent using a broadcast.	Uses a multicast address for updates.
Uses a metric based on how distant the remote network is to the router.	Is capable of using a complex metric.
Routing information learned from its neighbors.	Routing information learned from every router in the area.
The routing table is viewed from the perspective of each router.	The topological database is the same for every router in the area.
Uses Bellman Ford algorithm.	Uses the Dijkstra algorithm.
Does not consume many router resources, but is heavy in the use of network resources.	Uses many router resources, but is relatively low in its demand for network resources.
Maintains one domain in which all the routes are known.	Has a hierarchical design of areas that allow for summarization and growth.
Restricted by classful addressing scheme.	For effective use, the addressing scheme should reflect the hierarchical design of the network.
Involves slower convergence.	Involves quicker convergence.

Part II: OSPF

Chapter 6 Using OSPF in a Single Area

Chapter 7 Configuring OSPF in a Single Area

Chapter 8 Using OSPF Across Multiple Areas

Chapter 9 Configuring OSPF Across Multiple Areas

Part II covers the following Cisco BSCI exam topics:

- Describe the features and operation of single area OSPF
- Describe the features and operation of multiarea OSPF
- Given an addressing scheme and other laboratory parameters, identify the steps to configure a single-area OSPF environment and verify proper operation (within described guidelines) of your routers
- Given an addressing scheme and other laboratory parameters, identify the steps to configure a multiple-area OSPF environment and verify proper operation (within described guidelines) of your routers
- Identify the steps to verify OSPF operation in a single area
- Identify the steps to verify OSPF operation in multiple areas
- Interpret the output of various **show** and **debug** commands to determine the cause of route selection errors and configuration problems



This chapter covers the following topics, which you need to understand to pass the CCNP/CCDP/CCIP BSCI exam:

- Understanding OSPF fundamentals
- OSPF features
- OSPF operation in a single area
- OSPF network topologies

Using OSPF in a Single Area

The topics in this chapter detail the routing protocol OSPF. This chapter assumes knowledge of the previous chapter, which dealt conceptually with link-state routing protocols.

This chapter covers the essence of OSPF. It introduces OSPF by considering the protocol in its simplest form, within a single area. The basic operations of the protocol are explained in this chapter. Chapter 7, “Configuring OSPF in a Single Area,” builds on this chapter and explains how to configure, verify, and troubleshoot OSPF. Chapter 8, “Using OSPF Across Multiple Areas,” and Chapter 9, “Configuring OSPF Across Multiple Areas,” build further on this understanding and explain how OSPF works within a large multiarea network.

The topics in this chapter directly reflect questions on the exam. OSPF is the industry-standard interior routing protocol designed for use in large networks. Therefore, it is an obligatory subject in an exam on IP routing protocols. The following chapter assumes your comprehension of the subjects covered within this chapter.

“Do I Know This Already?” Quiz

The purpose of the “Do I Know This Already?” quiz is to help you decide what parts of this chapter to use. If you already intend to read the entire chapter, you do not necessarily need to answer these questions now.

The 12-question quiz, derived from the major sections in the “Foundation Topics” portion of the chapter, helps you to determine how to spend your limited study time.

Table 6-1 outlines the major topics discussed in this chapter and the “Do I Know This Already?” quiz questions that correspond to those topics.

Table 6-1 “Do I Know This Already?” Foundation Topics Section-to-Question Mapping

Foundation Topics Section	Questions Covered in This Section
Understanding OSPF Fundamentals	1–3
OSPF Features	4–6

(continues)

Table 6-1 “Do I Know This Already?” Foundation Topics Section-to-Question Mapping (Continued)

Foundation Topics Section	Questions Covered in This Section
OSPF Operation in a Single Area	7–9
OSPF Network Topologies	10–12

NOTE The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark this question wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

1. What is a neighbor in OSPF?
 - a. A neighbor is a router in the same area.
 - b. A neighbor is a router in the same classful network.
 - c. A neighbor is a router on a multiaccess link, with an adjacency with a DR.
 - d. A neighbor is another router with the same network address.
2. What is an adjacency in OSPF?
 - a. An adjacency is when another router has received an LSA from another area. The areas are adjacent.
 - b. An adjacency is the state that two neighbors can achieve after they have synchronized their OSPF databases.
 - c. An adjacent router is one that is sent a hello packet.
 - d. Routers connected across a WAN but not directly connected, for example Frame Relay, are considered adjacent to each other.
3. What is a designated router?
 - a. The router responsible for maintaining the SPF tree for a totally stubby area
 - b. The router responsible for summarizing routes to another areas
 - c. A router responsible for making adjacencies with all routers on a multiaccess link and maintaining those adjacencies
 - d. The router responsible for forwarding all the traffic across the global Internet

4. How often, by default, does OSPF send out hello packets on a broadcast multiaccess link?
 - a. By default, OSPF sends out hello packets every 30 seconds on a broadcast network.
 - b. By default, OSPF sends out hello packets every 40 seconds on a broadcast network.
 - c. By default, OSPF sends out hello packets every 3.3 seconds on a broadcast network.
 - d. By default, OSPF sends out hello packets every 10 seconds on a broadcast network.
5. If a router has an OSPF priority set to 0, what does this indicate?
 - a. A router with the OSPF priority set to 0 is one that can participate in the election of a DR. It has the highest priority.
 - b. A router with the OSPF priority set to 0 is one that will switch OSPF packets before it does anything else.
 - c. A router with the OSPF priority set to 0 is one that cannot participate in the election of a DR. It can become neither a designated nor a BDR.
 - d. A router with the OSPF priority set to 0 is one that cannot participate in the election of a DR, but it can become a BDR.
6. When a router sends an LSA on a multiaccess link, to what is it sent?
 - a. Designated router
 - b. Designated router and the BDR
 - c. To all routers on the link; all routers maintain adjacencies, but it is only the DR that updates the rest of the network
 - d. DR who updates the BDR every 3.3 minutes
7. What does it mean when an interface shows that it is in the init state?
 - a. That an interface is coming online, determining the IP address and OSPF parameters
 - b. That a router coming online is waiting for a hello from a neighbor
 - c. That this is an point-to-multipoint interface and is waiting to connect to the WAN cloud
 - d. Seen only on broadcast links, it shows that the election of the DR is in progress
8. If the network is stable and sees no changes, how often will it send LSAs? Why are these updates sent out periodically?
 - a. Every 30 minutes by default. This is to ensure the integrity of the topological databases.
 - b. Every 30 seconds by default. This is to ensure that the network is fully connected.

- c. Never, there is no need if the network is stable.
 - d. Whenever an LSA is received, this means there is a problem on the network that needs to be flooded through the network.
- 9. In learning a new route, what will OSPF do if a received LSA is not found in the topological database?
 - a. The LSA is flooded immediately out of all the OSPF interfaces, except the interface from which the LSA was received.
 - b. The LSA is dropped and a message is sent to the transmitting router.
 - c. The LSA is placed in the topological database and an acknowledgement is sent to the transmitting router.
 - d. The sequence numbers are checked, and if the LSA is valid, it is entered into the topology database.
- 10. What does NBMA stand for?
 - a. Nonbroadcast multiadjacencies.
 - b. Nonbroadcast multiaccess.
 - c. Nonbreachable multidigest.
 - d. Nonbackup multiarea.
- 11. Which of the following best describes a virtual link?
 - a. A dial-on-demand link that appears to the routing tables of OSPF as if it is always present, but is raised when needed
 - b. A connection to another autonomous system that simulates one autonomous system
 - c. A virtual connection to a remote area that does not have any connections to the backbone (Area 0)
 - d. Point-to-point and point-to-multipoint link across an NBMA cloud
- 12. RFC 2328 describes the operation of OSPF in two modes across an NBMA cloud. What are they?
 - a. Point-to-point and broadcast operation
 - b. Nonbroadcast multiaccess and broadcast operation
 - c. Point-to-point and point-to-multipoint operation
 - d. Nonbroadcast multiaccess and point-to-multipoint operation

The answers to this quiz are found in Appendix A, “Answers to Chapter ‘Do I Know This Already?’ Quizzes and Q&A Sections.” The suggested choices for your next step are as follows:

- **6 or less overall score** —Read the entire chapter. This includes the “Foundation Topics” and “Foundation Summary” sections, the “Q&A” section, and the “Scenarios” at the end of the chapter.
- **7–9 overall score** —Begin with the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. If you have trouble with these exercises, read the appropriate sections in “Foundation Topics.”
- **10 or more overall score** —If you want more review on these topics, skip to the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. Otherwise, move to the next chapter.

Foundation Topics

Understanding OSPF Fundamentals

OSPF stands for Open Shortest Path First, an open standard using the SPF algorithm, making it a link-state routing protocol. OSPF is an open standard because it was built by a standards committee. The term *open standard* means that anyone can read the rules or standard and write an application. The routing protocol as such belongs to no one vendor, but to everyone. This documentation is freely available, allowing OSPF to be developed and offered by every vendor. With the specifications in place, it is an obvious solution to connect various technologies and vendor solutions.

OSPF’s purpose as a routing protocol is to convey routing information to every router within the organizational network. The technology that has been selected is a link-state technology, which was designed to be very efficient in the way it propagates updates, allowing the network to grow or scale.

OSPF Terminology

OSPF is a sophisticated protocol, but it is in essence quite straightforward. As with a 19th century Russian novel, when you know the different names of the protagonists and how they interrelate, the rest is simple.

Table 6-2 explains briefly the OSPF terminology that you will see in the next few chapters.

Table 6-2 *OSPF Terms*

Term	Description
Adjacency	Formed when two neighboring routers have exchanged information and have the same topology table. The databases are synchronized, and they both see the same networks.
Area	A group of routers that share the same area ID. Each router in the area has the same topology table. Each router in the area is an internal router. The area is defined on an interface basis in the configuration of OSPF.
Autonomous system	Routers that share the same routing protocol within the organization.
Backup designated router (BDR)	The backup to the designated router (DR), in case the DR fails. The BDR performs none of the DR functions while the DR is operating correctly.

Table 6-2 OSPF Terms (Continued)

Term	Description
Cost	The metric for OSPF. It is not defined in the standard with a value. Cisco use the default of the inverse of bandwidth so that the higher the speed of the link, the lower the cost, and, therefore, the more attractive the path. This default can be overridden by a manual configuration. This should be done only if you have a full knowledge of the network.
Database descriptor	Referred to as DBDs or database descriptor packets (DDPs). These are packets exchanged between neighbors during the exchange state. The DDPs contain partial LSAs, which summarize the links of every router in the neighbor's topology table.
Designated router (DR)	Router responsible for making adjacencies with all neighbors on a multiaccess network, such as Ethernet or FDDI. The DR represents the multiaccess network, in that it ensures that every router on the link has the same topology database.
Dijkstra algorithm	A complex algorithm used by routers running link-state routing protocols to find the shortest path to the destination.
Exchange state	State in which two neighboring routers discover the map of the network. When these routers become adjacent, they must first exchange DDPs to ensure that they have the same topology table.
Exstart state	State in which the neighboring routers determine the sequence number of the DDPs and establish the master/slave relationship.
Flood	A term that refers to network information. When network information is flooded, it is sent to every network device in the domain.
Fully adjacent	When the routing tables of the two neighbors are fully synchronized, with exactly the same view of the network.
Init state	State in which a hello packet has been sent from the router, which is waiting for a reply to establish two-way communication.
Internal router	A router that has all its interfaces in the same area.
Link-state advertisement (LSA)	A packet describing a router's links and the state of those links. There are different types of LSAs to describe the different types of links. These are discussed in Chapter 9, "Configuring OSPF Across Multiple Areas."
Link-state database	Otherwise known as the topology map, the link-state database has a map of every router, its links, and the state of the links. It also has a map of every network and every path to each network.

(continues)

Table 6-2 OSPF Terms (Continued)

Term	Description
Link-state request (LSR)	When the router receives a DDP complete with a partial LSA, it compares the summarized information against the topological database. If either the LSA entry is not present or the entry is older than the DDP, it will request further information.
Link-state update (LSU)	Update sent in response to the LSR. It is the LSA that was requested.
Loading state	State in which, if the receiving router requires more information during the process in which two routers are creating an adjacency, it will request that particular link in more detail using the LSR packet. The LSR will prompt the master router to send the LSU packet. This is the same as an LSA used to flood the network with routing information. While the receiving router is awaiting the LSUs from its neighbor, it is in the loading state.
Neighbor	A router on the same link with whom routing information is exchanged.
Neighbor table	A table built from the hello messages received from the neighbors. The hello message also carries a list of the neighbors.
Priority	A Cisco tool by which the DR can be manually elected or, conversely, prevented from taking part in the DR/BDR election.
Shortest Path First (SPF)	The same as the Dijkstra algorithm, which is the algorithm used to find the shortest path.
SPF tree	A tree of the topological network. It can be drawn after the SPF algorithm has been run. The algorithm prunes the database of alternative paths and creates a loop-free shortest path to all networks. The router is at the root of the network, which is perceived from its perspective.
Topology table	The same as a link-state database. The table contains every link in the wider network.
Two-way state	State during the process in which two routers are creating an adjacency. The new router sees its own router ID in the list of neighbors, and a neighbor relationship is established. This is the stage before routing information is exchanged.

OSPF Features

OSPF has many features, the most important of which are dealt with in the following section in the context of the simplest OSPF network design, that of a single area. The concept of neighbors;

adjacent neighbors; DRs; and the role of the hello packet—which creates and maintains these neighbors, adjacencies, and DRs—are all considered in this section.

OSPF Neighbors

A neighbor in OSPF is a router that shares the same network link or the same physical segment. A router running OSPF discovers its neighbors by sending and receiving a simple protocol called the *Hello protocol*.

A router configured for OSPF sends out a small hello packet periodically (10 seconds is the default on broadcast multiaccess media). It has a source address of the router and a multicast destination address set to AllSPFRouters (224.0.0.5). All routers running OSPF (or the SPF algorithm) listen to the protocol and send their own hello packets periodically.

Adjacent OSPF Neighbors

After neighbors have been established by means of the Hello protocol, they exchange routing updates. This information about the network is entered into a database, called the topology table. From this database, the best paths to destinations are calculated and entered into the routing table. Therefore, the neighbor relationship is the key to understanding OSPF, as a router's neighbor gathers information about the network and passes it on to its directly connected neighbors.

When the topology databases of the neighbors are the same (synchronized), the neighbors are *fully adjacent*. To ensure that the link is maintained and the topology databases are up to date and accurate, the Hello protocol continues to transmit.

The transmitting router and its networks reside in the topology database for as long as the other routers receive the Hello protocol.

Advantages of Having Neighbors

There are obvious advantages to creating neighbor relationships. These advantages include the following:

- It is a mechanism for determining that a router has gone down (obvious because its neighbor no longer sends hello packets).
- Streamlined communication results because after the topological databases are synchronized, incremental updates will be sent to the neighbors as soon as a change is perceived, as well as every 30 minutes.
- Adjacencies created between neighbors control the distribution of the routing protocol packets.

The use of adjacencies and a neighbor relationship results in a much faster convergence of the network than can be achieved by RIPv1. This is because RIPv1 must wait for incremental updates

and holddown timers to expire on each router before the update is sent out. Convergence on a RIPv1 network can take many minutes, and the real problem is the confusion created by the different routing tables held on different routers during this time. This problem can result in routing loops and “black holes” in the network.

The DR

If routers are connected to a broadcast segment, one router on the segment is assigned the duty of maintaining adjacencies with all the routers on the segment. This router is known as the *designated router (DR)* and is elected by the use of the Hello protocol. The hello packet carries the information that determines the DR and the BDR, which you will learn more about in the next section, “BDRs.” The election is determined by either the highest IP address or the following command (if it is defined):

```
Router(config-if)#ip ospf priority number
```

The number in the **priority** command can be set between 0–255, where the higher the number, the greater the likelihood that this router will be selected as the DR.

All other routers peer with the DR, which informs them of any changes on the segment.

DRs are created on multiaccess links, because if there are many routers on the same segment, the intermesh of neighbor relationships becomes complex. Mathematically speaking, the number of adjacencies required for a full mesh is $n(n-1)/2$ and for a DR/BDR situation is $2n-2$.

On an FDDI ring, which forms the campus or building backbone, each router must form an adjacency with every other router on the segment. Although the Hello protocol is not networking-intensive, maintaining the relationships requires additional CPU cycles. Also, there is a sharp increase in the number of LSAs generated.

If one router is elected foreman of the link, responsible for maintaining adjacencies and forwarding updates, this dramatically reduces the overhead on the network.

BDRs

A network administrator does not want the responsibility of the segment to fall to one router, which poses the frightening situation of a single point of failure. Instead, you need to build redundancy into the network with the BDR. The BDR knows all the links for the segment. All routers have an adjacency not only with the DR, but also with the BDR, which in turn has an adjacency with the DR. If the DR fails, the BDR immediately becomes the new DR.

Electing the DRs and BDRs

You can manually elect the DRs and BDRs, or you can rely on the Hello protocol to select them dynamically, as described in the next sections.

Dynamic Election of the DR

When selected dynamically, the DR is elected arbitrarily. The selection is made on the basis of the highest router ID or IP address present on the network segment. Be aware that the highest IP address is the numerically highest number, not the class ranking of the addresses. Therefore, an elderly 2500 router with a Class C address of 192.168.250.249 might end up as a DR although there is a 7500 available on the segment that connects to the other segments. Unfortunately, the address of 10.10.10.1 is not as high as an old, frail, low-capacity router. This might not be the optimal choice.

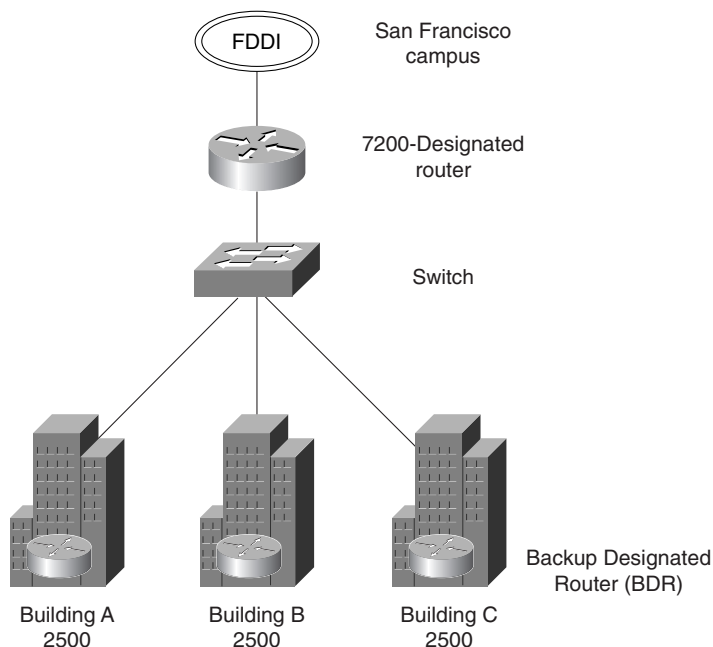
After the DR and BDR have been elected, all routers on the broadcast medium will communicate directly with the DRs. They will use the multicast address to all DRs. The BDR will listen but will not respond; remember, the BDR is the understudy waiting in the wings. The DR will send out multicast messages if it receives any information pertinent to the connected routers for which it is responsible.

Manual Configuration of the DR

To determine manually which router will be the DR, it is necessary to set the priority of the router. A router interface can have a priority of 0 to 255. The value of 0 means that the router cannot be a DR or BDR; otherwise, the higher the priority, the more favorable the chances are of winning the election.

If there is more than one router on the segment with the same priority level, the election process picks the router with the highest router ID. The default priority on a Cisco router is 1.

In Figure 6-1, the 2500 router for Building A, which is connected to the San Francisco campus via a hub, would be a reasonable choice as the DR. Although it is small, size is not as important as fault tolerance in this situation.

Figure 6-1 *The DR*

Because there are not many other routers on the segment, the number of LSAs and adjacencies that this router would have to record is small.

The larger 7200 Cisco router, which connects the building routers to the campus backbone, acts as the centralized router; therefore, the 7200 Cisco router makes sense as the router in charge of the connectivity of the campus buildings, allowing another router on the FDDI ring (not pictured) to take the DR responsibility for the FDDI segment. It would be a mistake to make the 7200 the DR for both networks, because this would increase the demand for resources and also would centralize all the responsibility on one router.

The Election of the DR

The following is the process used to elect the designated and BDRs:

All the neighbors who have a priority greater than 0 are listed.

1. The neighbor with the highest priority is elected as the BDR.
2. If there is no DR, the BDR is promoted as DR.

3. From the remaining routers, the router with the highest priority is elected as the BDR.
4. If the priority has not been configured, there will be a tie, because the default is to set the priority to 1.
5. If there is a tie because the priority has not been configured, the highest router IDs are used.

The Hello Packet

Although the routers running OSPF transmit a small packet, called the hello packet, to establish neighbor relations, it serves other functions. The various fields in the hello packet have specific responsibilities. Figure 6-2 shows the format of the hello packet. Table 6-3 describes each field.

Figure 6-2 *The Hello Packet*

Version #	1	Packet Length
Router ID		
Area ID		
Checksum		AuType
Authentication		
Authentication		
Network Mask		
HelloInterval	Options	Rtr Pri
RouterDeadInterval		
Designated Router		
Backup Designated Router		
Neighbor		
...		

Table 6-3 *The Hello Packet Fields*

Field	Characteristics	Function
Common OSPF Header		
Version #	The version of OSPF, which is currently version 2	To ensure the versions of OSPF are compatible
Packet Type=1	This states the type of OSPF packet after the header.	The Type 1 header is the hello packet.

(continues)

Table 6-3 The Hello Packet Fields (Continued)

Field	Characteristics	Function
Packet Length	This is the length of the packet including the OSPF header.	This field is used to identify the packet length.
Router ID	This is a 32-bit number. The highest IP address on the router is used as the ID. If a loopback address is configured, this will be used, even if it is not the highest address. If there are multiple loopbacks, the highest address is chosen.	This field identifies the router within the autonomous system. It is the ID of the originating router.
Area ID	This is the area ID of the originating router's interface.	The hello packet must come from a router within the same area to be valid.
Checksum	A checksum on the entire OSPF packet excluding the authentication field.	This is used to ensure the integrity of the packet.
AU Type	States the type of authentication used	Ensures the same authentication is used between systems
Authentication	64 bit authentication	Used for security between systems
Hello Packet Format		
Network Mask	The network mask for the transmitting interface	The mask must match the mask on the receiving interface, ensuring that they share a segment and network.
Hello Interval Options RouterDeadInterval	Used on broadcast, multiaccess networks: Dead Interval=40 Hello=10 sec Used on nonbroadcast networks: Dead Interval=120 Hello = 30	Hello maintains the presence of the router in its neighbor's databases. It works like a keepalive. The dead interval is how long the router waits before it determines that a neighbor is unavailable because it has not heard a hello packet within the prescribed time, that is, four times the hello timer.

Table 6-3 *The Hello Packet Fields (Continued)*

Field	Characteristics	Function
Neighbor	The router ID of a neighbor is entered in the neighbor table when a two-way (bidirectional) communication is established within the RouterDeadInterval. The communication is established when the router sees itself listed as a neighbor in the hello packet generated by another router on the same physical segment.	A neighbor is another router with which updates will be exchanged to synchronize databases.
Rtr Pri	This is the router priority of the source router interface. The higher the priority, the higher the likelihood of the router being selected as a DR or BDR.	This field is used to select the DR and BDR manually.
Designated Router	This is the address of the existing DR.	This field is used to allow the router to create unicast traffic to the DR router.
Backup Designated Router	This is the address of the existing BDR.	This field is used to allow the router to create unicast traffic to the BDR router.
Authentication	This specifies the authentication type and information. If set, the password must match the password stated on the router.	This field is used as security.

OSPF Operation in a Single Area

OSPF operates as a classic link-state routing protocol. It uses topology tables as well as the SPF tree as the basis of the SPF algorithm. This algorithm, created by Edsger Wybe Dijkstra, creates an SPF tree from the topology table. After calculating the algorithm on the SPF tree, the forwarding table is created. The forwarding table is, in fact, the routing table by another name. This section considers and describes how routes are entered into and removed from the routing table.

Creating and Maintaining the OSPF Routing Table

As discussed in the section “OSPF Neighbors,” after a neighbor is discovered in OSPF, an adjacency is formed. It is important to understand how the neighbor adjacency is formed and, in this context, to understand the other messages that the routers receive.

Routing tables are built in two different ways. Either established databases have to adjust to a change in the network, or a new router has to create the topology and forwarding databases when it enters the network.

Different techniques are used for these different routing table requirements. Essentially, the difference between the two techniques is simple:

- If a new router connects to a network, it will find a neighbor using the Hello protocol and will *exchange* routing information.
- If a change occurs in an existing network, the router that sees the change will *flood* the area with the new routing information.

Both of these events must occur as stated because, although the new router must learn the network topology, its addition is a change to the rest of the network.

These two requirements for updating the routing table use different technologies and OSPF protocols. These technologies and protocols are often confused, so the next sections take a moment to distinguish them. Understanding the distinction makes the OSPF operation much clearer.

How OSPF Builds the Routing Table on a New Router

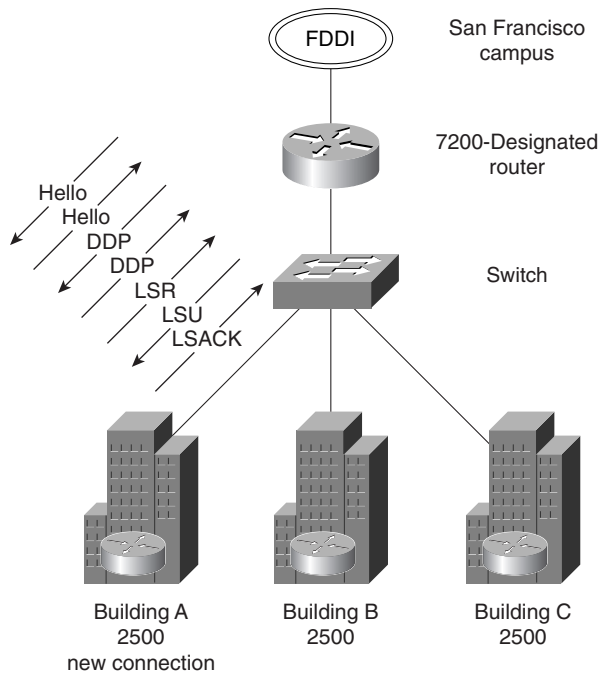
When a new router is added to the network, it builds a routing table by listening to the established routers with complete routing tables. Remember that every router within an area will have the same database and will know of every network within the area. The routing table built from this database is unique to the router because the decisions depend on the individual router's position within the area, relative to the remote destination network.

Five packet types are used to build the routing table for the first time:

- **Hello protocol** —Used to find neighbors and to determine the designated and BDR. The continued propagation of the Hello protocol maintains the transmitting router in the topology database of those that hear the message.
- **Database descriptor** —Used to send summary information to neighbors to synchronize topology databases.
- **LSR** —Works as a request for more detailed information, which is sent when the router receives a database descriptor that contains new information.
- **LSU** —Works as the LSA packet issued in response to the request for database information in the LSR packet. The different types of LSA are described in Chapter 8 in the section “Link-State Advertisements.”
- **Link-state acknowledgement** —Acknowledges the LSU.

Consider the case of a router joining the OSPF network for the first time. In Figure 6-3, the 2500 router in Building A at the San Francisco campus has just been connected.

Figure 6-3 *Joining an OSPF Network*



The next sections detail what happens when a router joins a network.

Finding Neighbors with the Exchange Process

When it is connected to the network and has been configured to run OSPF, the new router must learn the network from the systems that are up and running. The method described in this section is the same as for a stable network.

The process in this section shows the stages that the systems go through while exchanging information. You can see what stage an interface running OSPF is in with the command **show ip ospf neighbor** as well as the command **debug ip ospf adjacency**. Care should be taken with the **debug** command because it can be CPU-intensive.

The different stages or states that the router goes through while creating a neighbor relationship are shown in the following list:

1. **The down state** —The new router is in a *down state*. The 2500 router transmits its own hello packets to introduce itself to the segment and to find any other OSPF-configured routers. This is sent out as a hello to the multicast address 224.0.0.5 (AllSPFRouters). It sets the DR and BDR in the hello to be 0.0.0.0.
2. **The init state** —The new router waits for a reply. Typically this is four times the length of the hello timer. The router is in the *init state*. Within the wait time, the new router hears a hello from another router and learns the DR and the BDR. If there is no DR or BDR stated in the incoming hello, an election takes place. However, in accordance with the description of the Hello protocol, the DR has been elected: It is the 7200 router, which connects the campus to the campus backbone.

Upon hearing the Hello protocol from the 2500 router, a router on the segment adds the router ID of the 2500 and replies as a multicast (224.0.0.5) with its own ID and a list of any other neighbors.

3. **The two-way state** —The new router sees its own router ID in the list of neighbors, and a neighbor relationship is established. The new router changes its status to the *two-way state*.

Discovering Routes

The 2500 router and the designated router have now established a neighbor relationship and need to ensure that the 2500 has all the relevant information about the network. The 7200 router must update and synchronize the topology database of the 2500. This is achieved by using the exchange protocol with the database description packets.

The different stages or states that the router goes through while exchanging routing information with a neighbor are shown in the following list:

1. **The exstart state** —One of the routers will take seniority, becoming the master router. This is the *exstart state*. The two neighbors determine a master/slave relationship based on highest IP interface address. This designation is not significant; it just determines which router starts the communication.
2. **The exchange state** —Both routers will send out database description packets, changing the state to the *exchange state*.

In this example, the 2500 router has no knowledge and can inform the 7200 router only of the networks or links to which it is directly connected. The 7200 sends out a series of database description packets containing the networks held in the topology database. These networks are referred to as *links*.

Most of the information about the links has been received from other routers (via LSAs). The router ID refers to the source of the link information.

Each link will have an interface ID for the outgoing interface, a link ID, and a metric to state the value of the path. The database description packet will not contain all the necessary information, but just a summary (enough for the receiving router to determine whether more information is required or whether it already contains that entry in its database).

When the router has received the DDPs from the neighboring router, it compares the received network information with that in its topology table. In the case of a new router, such as the 2500, all the DDPs are new. Remember that the DDPs are simply a summary of the routes that the neighbor knows about. The different stages or states that the router goes through gathering routing information to update the topology database from a neighbor are shown in the following list:

3. **The loading state** — If the receiving router, the 2500, requires more information, it will request that particular link in more detail using the LSR packet. The LSR will prompt the master router to send the LSU packet. For example, if there is a discrepancy between the information in the received DDPs and the router's topology database, the router requests more detailed information from its neighbor about those routes of which it was unaware.

This process is the same as an LSA used to flood the network with routing information. While the 2500 is awaiting the LSUs from its neighbor, it is in the *loading state*.

4. **The full state** — When these LSRs are received and the databases are updated and synchronized, the neighbors are *fully adjacent*.

All the stages in updating a router's databases, as described in the numbered list, are illustrated in Figure 6-4.

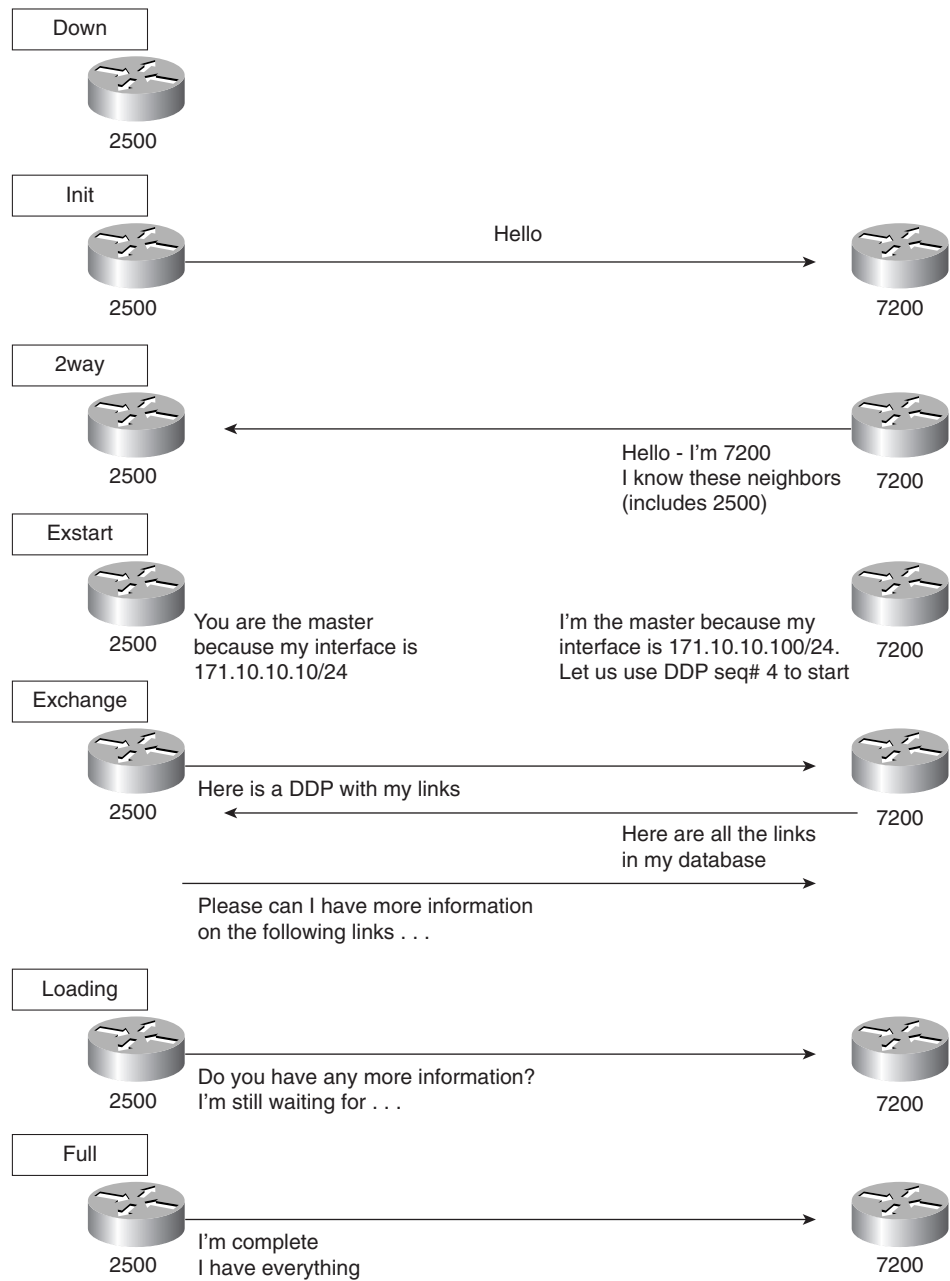
Now that you understand how OSPF learns about the connected network by forming adjacencies, the second stage is to learn how the neighbors flood information about their links throughout the network. The next section describes how the topology database, sometimes referred to as the link-state database, learns about the entire OSPF domain, or autonomous system.

The Topology Database

The topology database is the router's view of the network within the area. It includes every OSPF router within the area and all the connected networks. This database is indeed a routing table, but a routing table for which no path decisions have been made; it is at present a topology database.

The topology database is updated by the LSAs. Each router within the area has exactly the same topology database. All routers must have the same vision of the network; otherwise, confusion, routing loops, and loss of connectivity will result.

Figure 6-4 The Stages of Updating the Routers About the Network



The synchronization of the topology maps is ensured by the intricate use of sequence numbers in the LSA headers.

From the topology map, a routing database is constructed. This database will be unique to each router, which creates a routing database by running the *shortest path first* (SPF) algorithm called the *Dijkstra algorithm*. Each router uses this algorithm to determine the best path to each network and creates an *SPF tree* on which it places itself at the top, or root. If there are equal metrics for a remote network, OSPF includes all the paths and load balances the routed data traffic among them.

Occasionally, a link might flap or go up and down. This is more common on a serial line. If this happens, it could cause many LSAs to be generated when updating the network. To prevent this from happening, OSPF introduced timers. These timers force OSPF to wait before recalculating SPF. They are configurable.

NOTE Although RFC 2328 does not state the number of multiple, equal-cost paths that can be used at the same time, Cisco has defined a maximum of six paths that can be used simultaneously for load balancing.

Maintaining the Topological Database and the Routing Table

Now turn back to the 2500 router in Building A of the San Francisco campus in Figure 6-1. The router is now happily a member of the OSPF network. This section follows the process of hearing an update to the network in the form of an LSA.

As soon as a router realizes that there has been a change in the network topology, the router is responsible for informing the rest of the routers in the area. Typically, it will identify a change in the state of one of its links for one of the following reasons:

- The router loses the physical or data link layer connectivity on a connected network. The router propagates an LSU and sends it to the DR on a multiaccess network or the adjacent router in a point-to-point network. From there, it is flooded to the network.
- The router fails to hear either an OSPF Hello protocol or a data link Hello protocol. The router propagates an LSU and sends it to the DR on a multiaccess network or the adjacent router in a point-to-point network. From there, it is flooded to the network.
- The router receives an LSA update from an adjacent neighbor, informing it of the change in the network topology. The LSU is acknowledged and flooded out the other OSPF interfaces.

In any of these instances, the router will generate an LSA and flood it to all its neighbors.

This discussion now turns to the process initiated when a router receives such an update. For this purpose, return to the 2500 router connected to its DR, the 7200, in Figure 6-1.

Learning a New Route

When the 2500 router receives a network LSA update from the DR, it goes through the following logical steps:

1. The router takes the first entry from the update—the first network with information about the state of its link.
2. The router verifies that the type of LSA is one that can be accepted by this router.
3. Having ascertained that it is a valid LSA which it can receive, the router issues a lookup to its topological database.
4. If the LSA entry is *not* in the topological database, it is flooded immediately out all the OSPF interfaces, except for the receiving interface.
5. If the LSA entry is in the topological database, further questions are required.
6. The router determines whether the new LSA has a more recent (higher) sequence number.
7. If the sequence numbers are the same, the router calculates the checksum for the LSAs and uses the LSA with the higher checksum.
8. If the checksum numbers are the same, the router checks the MaxAge field to ascertain which is the most recent update.
9. Having found that the latest LSU is the one that was received, the router determines whether it has arrived outside the wait period, before another computation is allowed (minsLSarrival).
10. If the new LSA entry passes these tests, it is flooded out all the OSPF interfaces, except for the receiving interface.
11. The current copy replaces the old LSA entry. If there was no entry, the current copy is just placed in the database.
12. The received LSA is acknowledged.
13. If the LSA entry was in the database, but the LSA that has just been received has an older sequence number, the router asks whether the information in the database is the same.
14. If the information is the same and the new LSA has an older sequence number, the process discards the packet. It might be old news, but there is no inconsistency in the database.
15. If the information is different and the newly received LSA has an older sequence number, however, the receiving router discards the LSA update. It issues a copy of the LSA it has in its database, sending it out of the receiving interface to the source address of the out-of-date LSA. The logic is that the sending router has bad or old information and must be updated because its topological database is obviously not synchronized with the rest of the area.

This ensures that any packets that get out of sequence will be verified before action is taken. It also attempts to rectify a problem that it sees—that of multiple routers offering different paths because their topological databases are completely confused.

16. After the initial flood, things calm down, and updates are sent only when there are changes in the area or when the 30-minute timer goes off. This timer ensures that the databases stay synchronized.

This process shows some of the internal complexity of OSPF. As you can see, the internals are extremely detailed. Therefore, the design of any OSPF network should be very carefully thought out. The configuration of the routing protocol, on the other hand, is incredibly straightforward.

Choosing the Shortest Path First and Building the Routing Table

As with any routing protocol, OSPF examines all the available paths to every network that it knows about. It selects the shortest, most direct path to that destination. This section discusses the metric OSPF uses to select the shortest path and the routing table information needed after the shortest path is determined.

The Metric

As with all routing protocols, the decision as to which path to place into the routing table is based on the metric used by the routing protocol. RIP, for example, uses hop count, which shows how many routers must be passed through to get to the destination. When CPU and memory speeds were slower, the latency of traveling through the router had much higher implications on network performance. OSPF has few of those constraints and so chooses the metric of *cost*. Cost is not defined, however; it depends on the implementation of the protocol. The metric can be programmed to be either complex or simple. Cisco's implementation of a dynamic and default cost uses a predefined value based on the bandwidth of the router interface. The network administrator can manually override this default.

On occasion, the metric determines more than one path to the destination. These are known as multiple equal-cost paths.

The cost is applied to the outgoing interface. The routing process will select the lowest accumulated cost of the interfaces to the remote network.

If the network is manually configured, all routers connected to a particular network should agree on cost. Also, if manually configured, the cost should be thought through very carefully.

Information Needed in the Routing Table

Having determined the shortest path or multiple equal-cost paths, the routing process will need to supply additional information. To forward the data down the chosen path, the next logical hop, link,

and outgoing interface must be ascertained. The routing table, or *forwarding database*, as it is sometimes called, requires this information.

The operation of OSPF across WANs is slightly different. The next section considers OSPF used over these different technologies.

OSPF Network Topologies

How an OSPF protocol communicates via the Hello protocol to its neighbors depends on the physical medium being used. OSPF identifies five distinct network types or technologies:

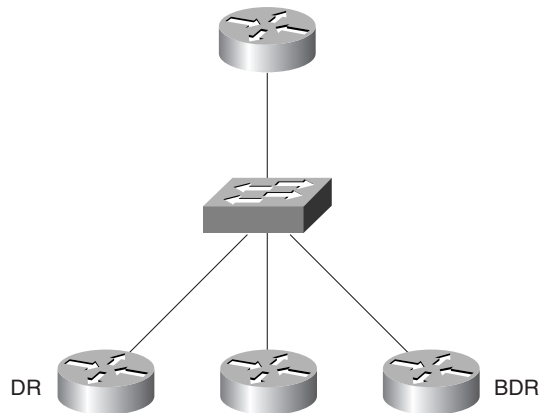
- Broadcast multiaccess
- Point-to-point
- Point-to-multipoint
- Nonbroadcast multiaccess (NBMA)
- Virtual links

The next sections describe each in more detail.

Broadcast Multiaccess Network

Broadcast multiaccess is any LAN network, such as Ethernet, Token Ring, or FDDI. In this environment, OSPF sends out multicast traffic. A DR and a BDR will be elected. Figure 6-5 illustrates a broadcast multiaccess network and the designated and BDRs.

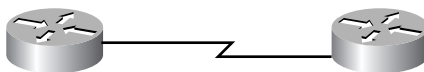
Figure 6-5 *A Broadcast Multiaccess Network*



Point-to-Point Network

Point-to-point technology is used where there is only one other router directly connected to the transmitting or receiving router. A typical example of this is a serial line. OSPF has no need for a DR or BDR in this scenario. OSPF messaging is sent using the multicast address for AllSPFRouters, 224.0.0.5. Figure 6-6 illustrates a point-to-point network.

Figure 6-6 *Point-to-Point Network*

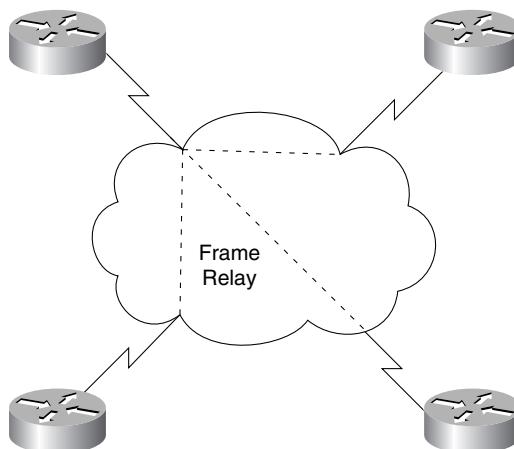


Point-to-Multipoint Network

Point-to-multipoint is a single interface that connects to multiple destinations. The underlying network treats the network as a series of point-to-point circuits. It replicates LSA packets for each circuit. OSPF traffic is sent as multicast. There is no DR or BDR election. This technology uses one IP subnet for all endpoints on the network.

Figure 6-7 illustrates a point-to-multipoint network.

Figure 6-7 *Point-to-Multipoint Network*



Nonbroadcast Multiaccess Network

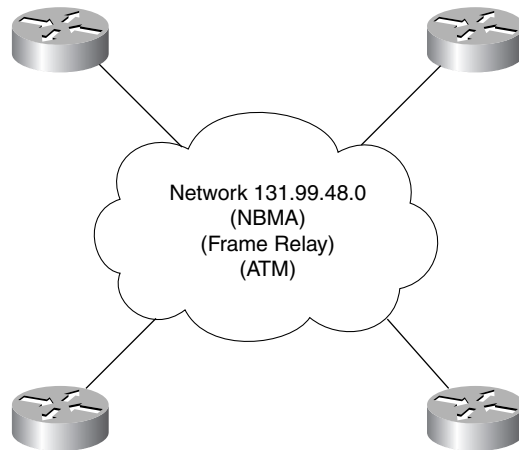
Physically, some point-to-multipoint networks cannot support multicast or broadcast traffic. In an NBMA topology, special configuration is required. NBMA physically resembles a point-to-point line, but in fact, many destinations are possible. WAN clouds, including X.25 and Frame Relay, are

examples of this technology. NBMA uses a fully meshed or partially meshed network. OSPF sees it as a broadcast network, and it will be represented by one IP subnet.

This technology requires manual configuration of the neighbors and the DR and BDR selection. The configuration options have increased with the different versions of Cisco IOS software.

DR and BDR routers are elected, and the DR will generate an LSA for the network. The DR and BDR must be directly connected to their neighbors. All network traffic sent between neighbors will be replicated for each physical circuit using unicast addresses, because multicast and broadcast addresses are not understood. Figure 6-8 illustrates an NBMA network.

Figure 6-8 *An NBMA Network*



Virtual Links

A virtual link is a virtual connection to a remote area that does not have any connections to the backbone (Area 0). Typically, this is because the network has become segmented. Although OSPF treats this link as a direct, single-hop connection to the backbone area, it is a virtual connection that tunnels through the network. The OSPF network traffic is sent in unicast datagrams across these links.

Having discussed OSPF network topologies, including WAN technologies, the next section, “OSPF Across NBMA Networks,” discusses the NBMA topologies in more detail. Remember that the method by which the routers in an OSPF network find one another and exchange information depends on the physical characteristics of the network.

OSPF Across NBMA Networks

An NBMA network has certain characteristics. The main ones are identified in the name of the technology: It is a network that cannot carry broadcast traffic but has multiple destinations. Examples of NBMA networks include Frame Relay, X.25, and ATM.

The crux of the problem is how OSPF operates using multicast traffic to exchange network information and to create adjacencies in order to synchronize databases across this WAN cloud without using the multicast addresses.

The solution to the problem varies, depending on the technology involved and the network design. The modes available fall into two technologies, within which there are additional options. The two technologies are point-to-point and NBMA.

The NBMA technology is then subdivided into two categories, under which different configuration options are available. These two categories are the RFC-compliant solution and the Cisco proprietary solution, as follows:

- **RFC-compliant** —The RFC-compliant category offers a standard solution, which is independent of the vendor platform. The configuration options are:
 - NBMA
 - Point-to-multipoint
- **Cisco-specific** —These configuration options are proprietary to Cisco and include:
 - Point-to-multipoint nonbroadcast
 - Broadcast
 - Point-to-point

The option you select depends on the network topology that is in use. The OSPF technology is separate from the physical configuration, and the choice of implementation is based on the design topology.

The Frame Relay topologies include:

- **Full mesh** —Every router is connected to every other router. This solution provides redundancy, and it might allow load sharing. This is the most expensive solution.
- **Partial mesh** —Some routers are connected directly; others are accessed through another router.
- **Star, or hub and spoke** —One router acts as the connection to every other router. This is the least expensive solution because it requires the fewest number of permanent virtual circuits (PVCs). A single interface is used to connect to multiple destinations.

Choosing a Topology

Some of the considerations in choosing the OSPF topology depend on its method of updating the network and its effect on network overhead. These considerations are mentioned in RFC 1586, which suggests that the different virtual circuits have different functions, as follows:

- **A point-to-point circuit** —Although no DR or BDR is required, each circuit will have an adjacency, which will create many more adjacencies on the network and will increase the need for network resources.
- **An NBMA environment** —This might require a DR and a BDR, unless the underlying technology is viewed as point-to point. This is economical for most routers, requiring only two adjacencies, except for the DR and BDR. However, it might require more administration in terms of configuration.

Subinterfaces

On a Cisco router, it is possible to configure a physical interface to be many logical interfaces. You can configure these subinterfaces to be point-to-point or point-to-multipoint. One of the main determining factors is the number of subnets to be used. A point-to-point interface requires its own subnet to identify it.

If you select the point-to-point option, managing the network is a little easier because the routers at each end create the adjacencies. The point-to-point option does require more network overhead and restricts some communication, in particular, the capability to indirectly connect through a hub router.

In a point-to-point network, the concept of a broadcast is not relevant because the communication is direct to another router. In a point-to-multipoint network, although OSPF simulates a broadcast environment, the network traffic is replicated and sent to each neighbor.

Table 6-4 indicates the characteristics and options for each case.

Table 6-4 OSPF over NBMA

	Point-to-Point Nonbroadcast	Point-to-Point	Broadcast	NBMA	Point-to Multipoint
Addressing	Unicast	Multicast	Multicast	Unicast	Multicast
DR/BDR	No	No	Yes	Manual Yes	No
Manual Configuration of Neighbors	Yes	No	No	Yes	No

Table 6-4 *OSPF over NBMA (Continued)*

	Point-to-Point Nonbroadcast	Point-to-Point	Broadcast	NBMA	Point-to- Multipoint
Hello	30 seconds Dead=120	10 seconds Dead=40	10 seconds Dead=40	30 seconds Dead=120	30 seconds Dead=120
RFC/Cisco	Cisco	Cisco	Cisco	RFC 2328	RFC 2328
Network Supported	Star Partial mesh	Star Partial mesh, using subinterfaces	Full mesh	Full mesh	Star Partial mesh (Seen as point- to-point)
Replicates Packets	Yes	Yes	Yes	Yes	Yes
Number of Subnets	1	Many (1 per circuit)	1	1	1

All the differing characteristics of the various network topologies can be very confusing, because it is not clear which type of network corresponds to a particular physical configuration. The following list clarifies the characteristics of the various network topologies:

- For serial interfaces with HDLC encapsulation, the default network type is point-to-point. Timers: hello 10, dead 40.
- For serial interfaces with Frame Relay encapsulation, the default network type is nonbroadcast. Timers: hello 30, dead 120.
- For serial interfaces with Frame Relay encapsulation and using point-to-point subinterfaces, the default network type is point-to-point. Timers: hello 10, dead 40.
- For serial interfaces with Frame Relay encapsulation and using point-to-multipoint subinterfaces, the default network type is nonbroadcast. Timers: hello 30, dead 120.

Now that you understand the mechanism of the OSPF routing protocol, this information will be useful in understanding how to configure the protocol on a Cisco router.

TIP If OSPF is used in an environment across different vendor equipment, it should be researched and tested to ensure interoperability.

Foundation Summary

The “Foundation Summary” section of each chapter lists the most important facts from the chapter. Although this section does not list every fact from the chapter that will be on your exam, a well-prepared candidate should, at a minimum, know all the details in each “Foundation Summary” before going to take the exam.

Table 6-5 explains common OSPF terms.

Table 6-5 *OSPF Terms*

Term	Description
Adjacency	Formed when two neighboring routers have exchanged information and have the same topology table. The databases are synchronized, and they both see the same networks.
Area	A group of routers that share the same area ID. Each router in the area has the same topology table. The area is defined on an interface basis in the configuration of OSPF.
Autonomous system	Routers that share the same routing protocol within the organization.
Backup designated router (BDR)	The backup to the designated router (DR), in case the DR fails. The BDR performs none of the DR functions while the DR is operating correctly.
Cost	The metric for OSPF. It is not defined in the standard with a value. Cisco use the default of the inverse of bandwidth so that the higher the speed of the link, the lower the cost, and, therefore, the more attractive the path. This default can be overridden by a manual configuration. This should be done only if you have a full knowledge of the network.
Database descriptor	Referred to as DBDs or database descriptor packets (DDPs). These are packets exchanged between neighbors during the exchange state. The DDPs contain a summary of the LSA, which describe the links of every router in the neighbor’s topology table.
Designated router (DR)	Router responsible for making adjacencies with all neighbors on a multiaccess network, such as Ethernet or FDDI. The DR represents the multiaccess network, in that it ensures that every router on the link has the same topology database.
Dijkstra algorithm	A complex algorithm used by routers running link-state routing protocols to find the shortest path to the destination.

Table 6-5 *OSPF Terms (Continued)*

Term	Description
Exchange state	Method by which two neighboring routers discover the map of the network. When these routers become adjacent, they must first exchange DDPs to ensure that they have the same topology table.
Exstart state	State in which the neighboring routers determine the sequence number of the DDPs and establish the master/slave relationship.
Flood	A term that refers to network information. When network information is flooded, it is sent to every network device in the domain.
Fully adjacent	When the routing tables of the two neighbors are fully synchronized, with exactly the same view of the network.
Init state	State in which a hello packet has been sent from the router, which is waiting for a reply to establish two-way communication.
Internal router	A router that has all its interfaces in the same area.
Link-state advertisement (LSA)	A packet describing a router's links and the state of those links. There are different types of LSAs to describe the different types of links.
Link-state database	Otherwise known as the topology map, the link-state database has a map of every router, its links, and the state of the links. It also has a map of every network and every path to each network.
Link-state request (LSR)	When the router receives a DDP complete with summary of the LSA information, it compares the LSA against the topological database. If either the LSA entry is not present or the entry is older than the DDP, it will request further information.
Link-state update (LSU)	Update sent in response to the LSR. It is the LSA that was requested.
Loading state	State in which, if the receiving router requires more information during the process in which two routers are creating an adjacency, it will request that particular link in more detail using the LSR packet. The LSR will prompt the master router to send the LSU packet. This is the same as an LSA used to flood the network with routing information. While the receiving router is awaiting the LSUs from its neighbor, it is in the loading state.
Neighbor	A router on the same link with whom routing information is exchanged.
Neighbor table	A table built from the hello messages received from the neighbors. The hello message also carries a list of the neighbors.
Priority	The means by which the DR can be manually elected—or, conversely, prevented from taking part in the DR/BDR election.

(continues)

Table 6-5 OSPF Terms (Continued)

Term	Description
Shortest Path First (SPF)	The same as the Dijkstra algorithm, which is the algorithm used to find the shortest path
SPF tree	A tree of the topological network. It can be drawn after the SPF algorithm has been run. The algorithm prunes the database of alternative paths and creates a loop-free shortest path to all networks. The router is at the root of the network, which is perceived from its perspective.
Topology table	The same as a link-state database. The table contains every link in the wider network.
Two-way state	State during the process in which two routers are creating an adjacency. The new router sees its own router ID in the list of neighbors, and a neighbor relationship is established. This is the stage before routing information is exchanged.

Table 6-6 describes the Hello packet.

Table 6-6 The Hello Packet

Field	Characteristics	Function
Router ID	This is a 32-bit number. The highest IP address on the router is used as the ID. If a loopback address is configured, this will be used, even if it is not the highest address. If there are multiple loopback addresses, the highest IP is chosen.	This field identifies the router within the autonomous system. It is the ID of the originating router.
Hello/Dead Intervals	Used on broadcast, multiaccess networks: Dead Interval=40 Hello=10 sec Used on nonbroadcast networks: Dead Interval=120 Hello = 30	Hello maintains the presence of the router in its neighbor's databases. It is a keepalive. The dead interval is how long the router waits before it determines that a neighbor is unavailable.

Table 6-6 *The Hello Packet (Continued)*

Field	Characteristics	Function
Neighbors	The router ID of a neighbor is entered in the neighbor table when a two-way (bidirectional) communication is established within the RouterDeadInterval. The communication is established when the router sees itself listed as a neighbor in the hello packet generated by another router on the same physical segment.	A neighbor is another router with which updates will be exchanged to synchronize databases.
Area ID	This is the area ID of the originating router's interface.	The hello packet must come from a router within the same area to be valid.
Router Priority	This is the priority of the source router interface. The higher the priority, the higher the likelihood of the router being selected as a DR or BDR.	This field is used to select the DR and the BDR manually.
DR IP address	This is the address of the existing DR.	This field is used to allow the router to create unicast traffic to the DR router.
BDR IP Address	This is the address of the existing BDR.	This field is used to allow the router to create unicast traffic to the BDR router.
Authentication Password	This is the authentication type and information. If this field is set, the password must match the password stated on the router.	This field is used as security.
Stub Area Flag	This field is set if the area is a stub area. All routers in the area must have this flag set.	This field identifies which type of LSAs will be transmitted and accepted.

Five packets are used to build the routing table for the first time:

- Hello protocol
- Database descriptor
- Link-state request
- Link-state update
- Link-state acknowledgement

Figure 6-9 is a flowchart illustrating the updating of the topological database

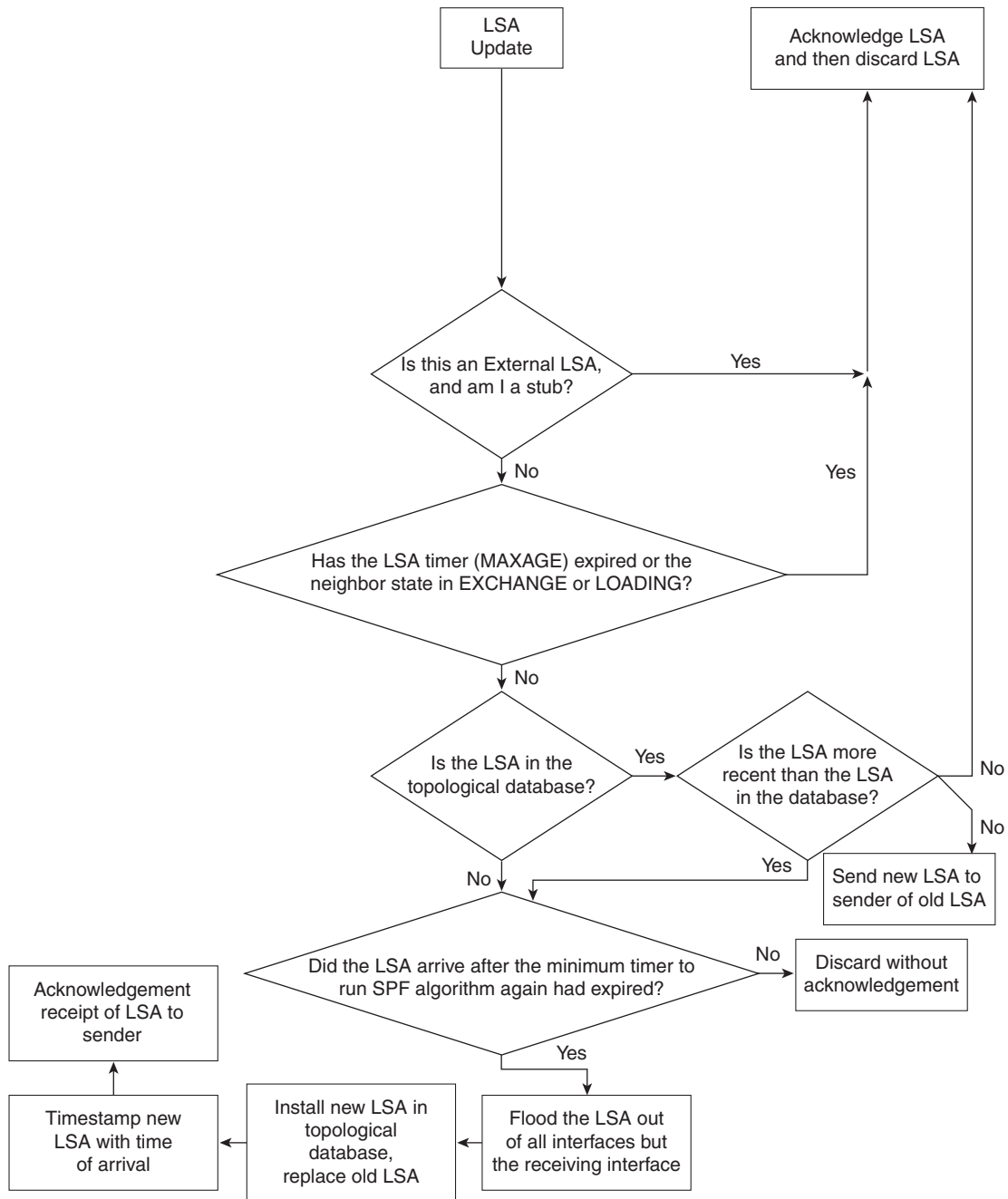
Figure 6-9 *Updating the Topological Database*

Table 6-7 indicates the characteristics and options for OSPF over NBMA.

Table 6-7 *OSPF over NBMA*

	Point-to-Multipoint Nonbroadcast	Point-to-Point	Broadcast	NBMA	Point-to- Multipoint
Addressing	Unicast	Multicast	Multicast	Unicast	Multicast
DR/BDR	No	No	Yes	Manual Yes	No
Manual Configuration of Neighbors	Yes	No	No	Yes	No
Hello	30 seconds Dead=120	10 seconds Dead=40	10 seconds Dead=40	30 seconds Dead=120	30 seconds Dead=120
RFC/Cisco	Cisco	Cisco	Cisco	RFC 2328	RFC 2328
Network Supported	Star Partial mesh	Star Partial mesh, using subinterfaces	Full mesh	Full mesh	Star Partial mesh (Seen as point-to-point)
Replicates Packets	Yes	Yes	Yes	Yes	Yes
Number of Subnets	1	Many (1 per circuit)	1	1	1

Q&A

As mentioned in the introduction, “All About the CCNP, CCDP, and CCIP Certifications,” you have two choices for review questions. The questions that follow next give you a bigger challenge than the exam itself by using an open-ended question format. By reviewing now with this more difficult question format, you can exercise your memory better and prove your conceptual and factual knowledge of this chapter. The answers to these questions are found in Appendix A.

For more practice with examlike question formats, including questions using a router simulator and multichoice questions, use the exam engine on the CD.

1. What information is held in the topology table?
2. What command is used to determine manually which router on a LAN will become the DR?
3. How many subnets are required in an OSPF configuration over a point-to-point network that has multiple connections?
4. State the different types of packets used to build a routing table for the first time.
5. In creating an adjacency, what is the exstart state?
6. What is the database descriptor and when is it used?
7. Explain the difference between an LSR and an LSA.
8. What packet is used to maintain the neighbor table?
9. What is the metric used by OSPF standards? Is this the same metric that Cisco uses?
10. Explain the meaning of the letters BDR.
11. What is used to elect the DR when the election is dynamic?
12. When a new router joins the OSPF network, will it learn about the rest of the OSPF network through the flooding method or the exchange method?
13. If an LSA is received that is present in the OSPF database, and the receiving LSA is older than the one currently held by the router, what action is taken?
14. A router has made a neighbor relationship with another router and exchanged DDP. Having compared the routing information from its neighbor, the router realizes that its topology database is incomplete. Name the different stages or states that a router goes through to update its topology database.
15. How many equal-cost paths will Cisco enter into the routing table?

16. An LSA is received by a router, and when checked against the topology database, it finds the LSA is new or a change in the status of an existing route that has been received. What action will the receiving router take?
17. Which NBMA configuration options are Cisco-specific?
18. What is the difference between a point-to-point interface and a point-to-multipoint interface?
19. What is the default network type for serial interfaces with HDLC encapsulation, and how often is the hello packet sent?
20. On a multiaccess link, what role does the BDR play?

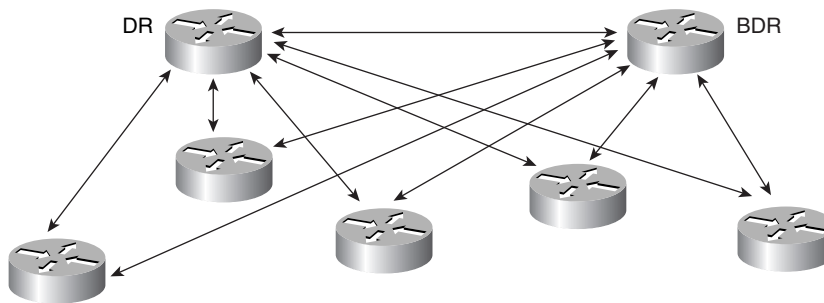
Scenarios

The following scenarios and questions are designed to draw together the content of the chapter and to exercise your understanding of the concepts. There is not necessarily a right answer. The thought process and practice in manipulating the concepts is the goal of this section. The answers to the scenario questions are found at the end of this chapter.

Scenario 6-1

In Figure 6-10, all routers share a common multiaccess segment. Because of the exchange of hello packets, one router is elected the DR and another is elected the BDR. Use Figure 6-10 to answer the following questions.

Figure 6-10 *Network Diagram for Scenario 6-1*



1. Which parameter determines the router that will be selected as the DR in an OSPF network?
2. Could a router with a priority value of zero assume the role of a DR or a BDR in the OSPF network shown in Figure 6-10?
3. How is the OSPF router ID determined on a Cisco router?
4. What is the role of the DR and BDR in the OSPF network shown in Figure 6-10?

Scenario Answers

The answers provided in this section are not necessarily the only possible answers to the questions. The questions are designed to test your knowledge and to give practical exercise in certain key areas. This section is intended to test and exercise skills and concepts detailed in the body of this chapter.

If your answer is different, ask yourself whether it follows the tenets explained in the answers provided. Your answer is correct not if it matches the solution provided in the book, but rather if it has included the principles of design laid out in the chapter.

In this way, the testing provided in these scenarios is deeper: It examines not only your knowledge, but also your understanding and ability to apply that knowledge to problems.

If you do not get the correct answer, refer back to the text and review the subject tested. Be certain to also review your notes on the question to ensure that you understand the principles of the subject.

Scenario 6-1 Answers

1. *Which parameter determines the router that will be selected as the DR in an OSPF network?*

The router with the highest OSPF priority on a segment will become the DR for that segment. The default for the interface OSPF priority is 1. If multiple routers have the same priority, the router with the highest RID will be selected as the DR.

2. *Could a router with a priority value of zero assume the role of a DR or BDR in the OSPF network shown in Figure 6-10?*

No. A priority value of zero indicates an interface is not to be elected as a DR or BDR. The state of the interface with priority zero will be DROTHER.

3. *How is the OSPF router ID determined on a Cisco router?*

The OSPF router ID is the highest IP address on the box, or the highest loopback address, if one exists.

4. *What is the role of the DR and the BDR in the OSPF network shown in Figure 6-10?*

Instead of each router exchanging updates with every other router on the segment, every router will exchange the information with the DR and the BDR. The DR and the BDR will relay the information to everybody else. In mathematical terms, the adjacencies required for a full mesh is $n(n-1)/2$ and for a DR/BDR situation is $2n-2$.



This chapter covers the following topics, which you need to understand to pass the CCNP/CCDP/CCIP BSCI exam:

- Configuring OSPF in a single area
- Configuring OSPF over an NBMA topology
- Checking the configuration of OSPF on a single router
- Troubleshooting OSPF in a single area

Configuring OSPF in a Single Area

This chapter explains how to configure, verify, and troubleshoot OSPF. This chapter assumes knowledge of the previous chapter, which dealt conceptually with the theory and operation of OSPF in a single area. This chapter covers how to configure OSPF in a single area, which is the simplest design. Although this configuration does not exploit the strengths of the link-state protocol, it introduces the fundamentals of OSPF configuration. You can build on this knowledge in the subsequent chapters that deal with the design and configuration of OSPF in a multiple-area environment.

This chapter assumes your comprehension of the subjects covered within Chapter 6, “Using OSPF in a Single Area.”

“Do I Know This Already?” Quiz

The purpose of the “Do I Know This Already?” quiz is to help you decide what parts of this chapter to use. If you already intend to read the entire chapter, you do not necessarily need to answer these questions now.

The 14-question quiz, derived from the major sections in the “Foundation Topics” portion of the chapter, helps you to determine how to spend your limited study time.

Table 7-1 outlines the major topics discussed in this chapter and the “Do I Know This Already?” quiz questions that correspond to those topics.

Table 7-1 “Do I Know This Already?” Foundation Topics Section-to-Question Mapping

Foundation Topics Section	Questions Covered in This Section
Configuring OSPF in a Single Area	1–6
Configuring OSPF over an NBMA Topology	7–9
Checking the Configuration of OSPF on a Single Router	10–12
Troubleshooting OSPF in a Single Area	13–14

NOTE The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark this question wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

1. At what level is the area for OSPF defined?
 - a. Network
 - b. Router
 - c. Process
 - d. Interface
2. Which of the following best define the purpose of the **network** command?
 - a. To state the networks that are to be summarized
 - b. To state the networks to be advertised
 - c. To define the OSPF areas and associate the interface with the area
 - d. To identify the interfaces through which OSPF packets are to be sent and received
3. Which of the following are valid **network** commands?
 - a. network 10.10.0.0 255.255.0.0 area 0
 - b. network 10.10.0.0 0.0.255.255 area 0.0.0.5
 - c. network 10.10.0.0 255.255.0.0 area 5
 - d. network 10.10.0.0 0.0.255.255 area 0.0.1.10
4. The default selection of the router ID would select a router ID by which of the following?
 - a. Highest IP address
 - b. Highest loopback address
 - c. No defaults; must be manually configured
 - d. Highest area ID, changing it to a dotted decimal notation
5. Which interface setting affects the default OSPF metric that is being run on a Cisco system?
 - a. There is no default metric setting; it is an industry standard and must be manually configured.
 - b. The error rate on the interface changes the cost of the link.

- c. The bandwidth setting changes the OSPF metric by default on Cisco routers.
 - d. Changing the **priority** command changes the OSPF process, which affects the cost of the link.
6. Which best describes a loopback interface in OSPF?
- a. The loopback interface is used for testing the router.
 - b. The loopback interface is a virtual interface that exists only in software.
 - c. The loopback interface is an interface that has been terminated for troubleshooting.
 - d. A loopback interface is an interface on FDDI.
7. When is the **neighbor** command used in configuring OSPF over a nonbroadcast NBMA link?
- a. Neighbors are defined to allow the manual configuration of the neighbor table.
 - b. Nonbroadcast multiaccess ATM clouds need neighbors to be defined so that the address can be appended to each cell.
 - c. In a nonbroadcast point-to-multipoint network, the **neighbor** command is used to define the next hop in OSPF.
 - d. In a nonbroadcast environment, the DR and BDR must be configured with a static list of the other routers attached to the cloud so that they can become neighbors and create adjacencies.
8. Where would you configure a point-to-point interface for OSPF over an NBMA topology?
- a. At the interface level
 - b. At the subinterface
 - c. Under the routing process
 - d. Under the subinterface
9. You use the broadcast mode to avoid using the **neighbor** command and all the attendant configurations. Which of the following commands is correct?
- a. **ip ospf ptmp**
 - b. **ip ospf point to point**
 - c. **ip ospf network broadcast**
 - d. **ip ospf broadcast**

10. Which OSPF database is shown with the **show ip ospf database** command?
 - a. Forwarding database
 - b. Topology database
 - c. Neighbor database
 - d. SPF tree

11. Which of the following are shown in the command **show ip ospf interface** ?
 - a. Link State Update Interval is 00:30:00
 - b. Network Type broadcast
 - c. Transmit delay is 1 sec
 - d. Dead timer due in 00:00:34

12. Which of the following commands shows the DR?
 - a. **show ip ospf neighbor detail**
 - b. **show ip ospf**
 - c. **show ip ospf database**
 - d. **show ip ospf interface**

13. Which packets are shown in the command **debug ip packet** ?
 - a. Discarded
 - b. Received
 - c. Generated
 - d. Forwarded

14. Which **debug** command shows the changes in adjacencies, flooding information, designated router selection, and shortest path first (SPF) calculations?
 - a. **debug ip packets**
 - b. **debug ip ospf events**
 - c. **debug ospf events**
 - d. **debug ospf packets**

The answers to this quiz are found in Appendix A, “Answers to Chapter ‘Do I Know This Already?’ Quizzes and Q&A Sections.” The suggested choices for your next step are as follows:

- **8 or less overall score—** Read the entire chapter. This includes the “Foundation Topics” and “Foundation Summary” sections, the “Q&A” section, and the “Scenarios” at the end of the chapter.
- **9–11 overall score—** Begin with the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. If you have trouble with these exercises, read the appropriate sections in “Foundation Topics.”
- **12 or more overall score—** If you want more review on these topics, skip to the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. Otherwise, move to the next chapter.

Foundation Topics

Configuring OSPF in a Single Area

When configuring any device, it is important to establish why you are configuring the system and what you are trying to achieve.

This section examines the configuration of a Cisco router for OSPF within a single area. The commands are few and simple; the implications are somewhat more difficult.

This section covers the following:

- Configuration of OSPF
 - Required configuration
 - Optional configuration
- Commands
 - What each configuration command achieves
 - How the configuration command achieves its goal

Required Commands for Configuring OSPF on an Internal Router

In this chapter, you learn to configure an internal router within a single area. An internal router is one that is within an area and whose sole function for OSPF is to route traffic within the area.

The router needs to understand how to participate in the OSPF network. Therefore, it requires the following:

- **The OSPF process** — The routing protocol needs to be started on the router.
- **Participating router interfaces** — The router might not want to have all its interfaces send or receive OSPF routing updates. A classic example is a dialup line to a remote office. If there is only one subnet at the remote office, it would be more efficient to use default and static route commands, because any updates would dial the line.
- **Identification of the area** — The router defines which area it is in on a per-interface basis.

- **A router ID (RID)**—This allows the router to be identified by the other routers in the network. The algorithm used to create the routing table builds a graph from a single point. IP addresses are usually assigned to interfaces, so the router needs to assign an IP address to represent the router itself; this is the RID. The ID of the router advertising a link is used to determine the next logical hop, for example, if that link is used in the path selection to a remote network.

The following two commands are required for configuring OSPF on a single internal router:

- **router ospf** *process-number* command
- **OSPF network** command

Enabling the OSPF Routing Protocol

When configuring the router for the first time, there is no IP routing protocol running on the Cisco router (unless the SETUP script is used). This is not true of other protocols, however; for example, if an IPX network address is configured on an interface, the IPX RIP process will be automatically started.

To configure OSPF as the routing protocol, use the following command:

```
Router(config)#router ospf process-number
```

Here, *process-number* is a number local to the router. It is possible to have more than one process running on a router, although this is an unusual and expensive configuration in terms of router resources. Repeating the command with another ID number will create another process. One possible scenario for this configuration is a service provider that wants to separate its OSPF domain from its customer.

The process number does not have to be the same on every router in the area or the autonomous system. In the interest of sanity, however, many administrators make it the same number.

NOTE A common error in configuration is to confuse the process ID with the RID or the area ID. These are not related in any way. The process ID is simply a mechanism to allow more than one process to be configured on a router. The RID is the mechanism by which a router is identified within the OSPF domain, and the area ID is a mechanism of grouping routers that share full knowledge of OSPF-derived routes within the OSPF domain.

Enabling the OSPF network Command

Although you have turned on OSPF, it has no information about how to operate. The networks that are to participate in the OSPF updates, and the area that they reside in, must be defined. If the following information is not specified, the process will have nothing to do:

```
Router(config-router)#network network-number wildcard-mask area area-number
```

This command deserves a moment's explanation because it is the cause of many errors in configuration.

The **network** command in OSPF plays a similar role to that of the **network** command in RIP or IGRP. The difference is the level of granularity afforded to the administrator. In RIP and IGRP, the **network** command is defined at the class level. In OSPF, it is possible to identify the specific address of an interface.

The additional parameter *area* states the area that the interface inhabits. This allows a router to have different interfaces in different areas, making it an area border router (ABR). The *area-number* is a 32-bit field and the format can take one of two forms: The first is a simple decimal, and the second is a dotted decimal format similar to that of an IP address. Some implementations of OSPF might only understand one of the formats, but Cisco will accept either form.

However, it is important to remember that decimal and dotted decimal use different numbering systems. For example, while 0.0.0.5 and 5 are the same, the area 0.0.4.6 is equivalent to 1030 in decimal. The dotted decimal notation is a stream, which you should think of as a continuation of the binary numbers; it does not start again as in an IP address. So the decimal number 1030 is

1024+6

or

00000000.00000000.00000100.00000110

What the **network** Command Will Do

After the **network** command has been entered, OSPF identifies which interfaces are participating in OSPF by comparing the interface IP address with the address given in the **network** command, filtered through the wildcard mask. The wildcard mask states how much of the address to pay attention to. The wildcard mask could look at just the class of address, such as everything in network 10.0.0.0, for example. At the other extreme, the mask can be more specific and identify an interface address. All interfaces that match the given network number will reside in the area specified in the **network** command.

CAUTION Take great care in choosing the wildcard mask. Remember that it follows the same format as the wildcard mask in an access list. It is extremely easy to make errors in the configuration, and those errors might be difficult to find.

After identifying the interfaces on the router that are participating in the OSPF domain, the following happens.

1. Updates will be received on the interface.
2. Updates will be sent out of the interfaces.
3. The interface will be placed in the defined area.
4. If appropriate, the Hello protocol will be propagated. Depending on the interface type, a default hello and dead interval are defined based on the OSPF network type.

This **network** command has many of the same characteristics as an access list. The wildcard mask has the same format and enables you to group interfaces into an area. It follows the same top-down logic of a link list, as seen in an access list.

NOTE If there are stub networks connected to a router, it is useful to issue the command **redistribute connected subnets** . This command is issued as part of the router process configuration, and it includes the connected subnets in OSPF advertisements without actually running OSPF on them. This is very useful for real OSPF configurations, particularly those that involve WAN pay-per-packet, low-bandwidth links.

Configuration Examples

The following examples show how one command can cover all router interfaces, and also how each individual interface can be specified.

Given a router with six interfaces, three with addresses in the 10.0.0.0 class and three with addresses in the 172.16.0.0 class, the following would configure all interfaces to participate in OSPF area 0:

```
Router(router-config)#network 0.0.0.0 255.255.255.255 area 0
```

The following would have only the interfaces addressed from 10.0.0.0 participating in OSPF area 0:

```
Router(config-router)#network 10.0.0.0 0.255.255.255 area 0
```

The next example shows only two specific interfaces participating in OSPF area 0:

```
Router(config-router)#network 10.12.0.1 0.0.0.0 area 0
Router(config-router)#network 172.16.15.1 0.0.0.0 area 0
```

Why Is the network Command so Complex?

It is reasonable to ask why OSPF is so much more complex than either IGRP or RIP in this instance. The answer is that the level of precision available in the OSPF **network** command provides the capability to place different interfaces into different areas on the same router. The need for this complexity is not obvious in this example because an internal router is being configured within a single area.

The flexibility in defining which interfaces reside in which area is considered in Chapter 9, “Configuring OSPF Across Multiple Areas,” in the section “Required Configuration Commands for a Multiarea OSPF Network.”

Options for Configuring OSPF on an Internal Router

The following options are not necessary to make OSPF function properly within an area. However, they might be useful in your network design:

- The loopback interface
- The **cost** command
- The **priority** command
- The RID

The following sections describe each option in more detail.

The Loopback Interface and the Router ID

The router needs an ID to participate in the OSPF domain. The RID is used to identify the source of LSA updates as shown in the OSPF database. This ID takes the form of an IP address. The address can be either defined by the administrator or left to the whim of the router. Most people define the ID so that it is easier to track events in the network, for internal documentation, and for other system-administration purposes.

The use of *loopback interface* addresses is often used to define the RID, as described in the following section. A loopback interface is a virtual interface, which has the advantage of never going down because it has no physical characteristics.

The Default Router ID Selection

The most common method of defining the RID is to use the defaults offered by Cisco. The default RID is taken from the highest IP address assigned to a loopback interface. If no loopback is defined, then OSPF takes the highest IP interface address as the RID.

If no ID is stated, the router will take the highest IP address configured on a loopback interface. Although it is unlikely that this address will change, it is possible. From an administrative viewpoint, such a change would introduce an unnecessary level of chaos into the network.

Manual Configuration of Router ID

The command to define the OSPF RID is within the router configuration. If there is no RID defined, there are other methods. The Cisco rule states that the RID will be taken from the address of the

loopback interface. If no loopback interface is defined, it uses the highest IP address of the active interfaces configured on the router.

Once the RID of the router has been chosen, the RID is not dependent on whether the interface is active or even functional until the router is rebooted. At that point, a different RID would be chosen from the active IP addresses. This could break some OSPF configurations, such as virtual links. You are therefore advised to configure the loopback interface. Because a virtual interface does not exist physically, it can never go down. Therefore, the OSPF RID is not vulnerable to hardware interface problems if the router reboots.

It is possible to have multiple loopback interfaces, in which case the loopback interface with the highest IP address will be selected if no RID has been configured. Many organizations choose a different addressing scheme for the loopbacks to distinguish them easily when troubleshooting. Remember that each interface requires a separate subnet. The use of a private address from RFC 1918 might be wise. Private addresses will not deplete the IANA address that is being used by the organization and have the advantage of being easily distinguished for administrative documentation.

The following shows how to configure the RID:

```
Router(config)#router ospf
Router(config-router)# router-id ip-address
```

The following shows how to configure a loopback interface:

```
Router(config)# interface loopback interface-number
Router(config-if)# ip address ip-address subnet-mask
```

NOTE When designing a network, consider whether to include the loopback interface address in the **network** commands. There are both advantages and disadvantages to this, and they should be researched in any network design. If the organization is running out of valid addresses, it might be advisable to use the loopback address only as an RID and not to insert it into the routing table. The disadvantage of this configuration is that it cannot be pinged for testing. This is known as a bogus RID. The preferred configuration would be to have an address in the routing table. These addresses are assigned a /32 subnet mask.

Changing the Default Metric Using the **cost** Command

Another command that might be useful is the **cost** command. This command manually overrides the default cost that the router assigns to the interface. The default cost is calculated based on the bandwidth parameter assigned to the outgoing interface with the **bandwidth** command.

The **cost** command syntax is as follows:

```
Router(config-if)# ip ospf cost cost
```

A lower cost increases the likelihood that the interface will be selected as the best or shortest path. The range of values configurable for the cost of a link is 1 to 65535.

In general, the path cost in Cisco routers is calculated using the formula $10^8/\text{bandwidth}$. Table 7-2 shows examples of default costs.

Table 7-2 *Default Costs in OSPF*

Link Type	Default Cost
56-kbps serial link	1785
T1 (1.544-Mbps serial link)	64
Ethernet	10
16-Mbps Token Ring	6
FDDI	1

NOTE Serial lines have many different speeds. The default bandwidth is 1.544 Mbps. If the line is a slower speed, use the **bandwidth** command to specify the real link speed. The cost of the link will then change to correspond to the bandwidth that you configured.

As shown in Table 7-2, the calculation of bandwidth gives FDDI a metric of 1. If you have multiple links with high bandwidth, you might want to have a higher number than the default cost in order to differentiate the cost on those links.

It is also possible to control how OSPF calculates default metrics for the interface. Use the **ospf auto-cost reference-bandwidth** router global configuration command to change the numerator of the previous OSPF cost formula:

```
Router(config-router)# ospf auto-cost reference-bandwidth reference-bandwidth
```

Here, *reference-bandwidth* is in megabytes per second. The range is 1 to 4,294,967; the default is 100. This means that a cost of 1 = 100 Mbps, and a cost of 10 = 100/10 = 10 Mbps. If you are moving to gigabit, you would want 1000 Mbps = 1 Gbps.

Any change using the **ospf auto-cost reference-bandwidth** command should be done on all routers in the autonomous systems so that they all use the same formula to calculate cost. The value set by the **ip ospf cost** command overrides the cost resulting from the **auto-cost reference-bandwidth** command.

In some of the Cisco IOS software documentation, the **auto-cost** command is documented as **ospf auto-cost**. However, **auto-cost** is the actual command in the Cisco IOS. Check the command reference set for the command for your IOS version.

Considerations in using the **cost** command include the following:

- Never change defaults unless you can explain why the change is necessary. Reasons for using the **cost** option in OSPF include the following:
 - You want to maintain interoperability among different vendors running OSPF.
 - There is a design reason to choose a different path than the one selected by the Cisco default metric.
 - You want to allow greater granularity in the application of the cost metric.
- If you override the default by manual configuration, it is important that you consider the physical and logical topology map of the network. Any change to the metric might change the traffic patterns in the network.

Determining the Designated Router Using the **priority** Command

The last optional command to consider is the **priority** command. You use this command to determine the designated router (DR) and backup designated router (BDR) on a multiaccess link. Remember that the Hello protocol carries the priority field and is the mechanism by which the DR and BDR are elected. To be “up for election,” the priority must be a positive integer between 1 and 255. If the priority is 0, the router cannot participate in the election. The higher the priority, the greater the likelihood of being elected. If no priority is set, all Cisco routers have a default priority of 1, and the highest RID is always used as a tiebreaker.

Reasons for increasing the router priority include the following:

- The router has greater CPU and memory than the others do on the LAN.
- The router is the most reliable router on the segment.
- All the other routers on the LAN connect to stub networks. They all form the access layer of the network.
- There are point-to-multipoint connections in an NBMA cloud, and the hub router needs to be configured as the centralized resource, requiring it to be the DR.
- The router is an ABR, and you do not want it to consume more resources as a DR, so another router on the subnet either has its priority increased or the ABR has its priority decreased.

The following section shows these commands in context to make their use and functionality much more apparent.

A Working Configuration of OSPF on a Single Router

Example 7-1 is a working configuration of OSPF on a single router. Use this example in conjunction with Figure 7-1.

The San Jose router is selected as the DR, after its priority is set to 100, and the cost of the fast Ethernet interface is set to 10, overriding the default cost.

Example 7-1 *Configuring OSPF*

```
SanJose(config)#router ospf 100
SanJose(config-router)#network 140.100.0.0 0.0.255.255 area 3
SanJose(config-router)#interface FastEthernet1/0
SanJose(config-if)#ip address 140.100.17.129 255.255.255.240
SanJose(config-if)#ip ospf priority 100
SanJose(config-if)#no shutdown
SanJose(config-if)#interface FastEthernet3/0
SanJose(config-if)#ip address 140.100.17.193 255.255.255.240
SanJose(config-if)#ip ospf cost 10
SanJose(config-if)#no shutdown
SanJose(config-if)#interface Fddi2/0
SanJose(config-if)#ip address 140.100.32.10 255.255.255.240
SanJose(config-if)#no ip directed-broadcast
SanJose(config-if)#no shutdown
```

Configuring OSPF over an NBMA Topology

The design considerations of running OSPF over an NBMA topology require thought and research because the configuration choices that are made will dramatically affect your network.

This section covers one of the common choices on a partially meshed network, illustrates the kind of decisions that need to be made, and shows the configuration that would be implemented.

If the network is partially meshed, then the choice to use only point-to-point subnets can waste addresses. If you choose a point-to-multipoint configuration, the network uses one subnet, and there is no DR/BDR negotiation. This configuration has the advantage of saving addresses but behaves as if it is a series of point-to-point links. If private addressing is used, this might not be a consideration.

When the decision is made as to which technology is to be implemented, whether you are using point-to-point or point-to-multipoint, the configuration is straightforward. The choice is defined on the interface as an **ip ospf network** command. The **network** command syntax is as follows:

```
Router(config-if)#ip ospf network {broadcast | non-broadcast | {point-to-multipoint
[non-broadcast]}}
```

Figure 7-1 Diagram for the Configuring OSPF Example

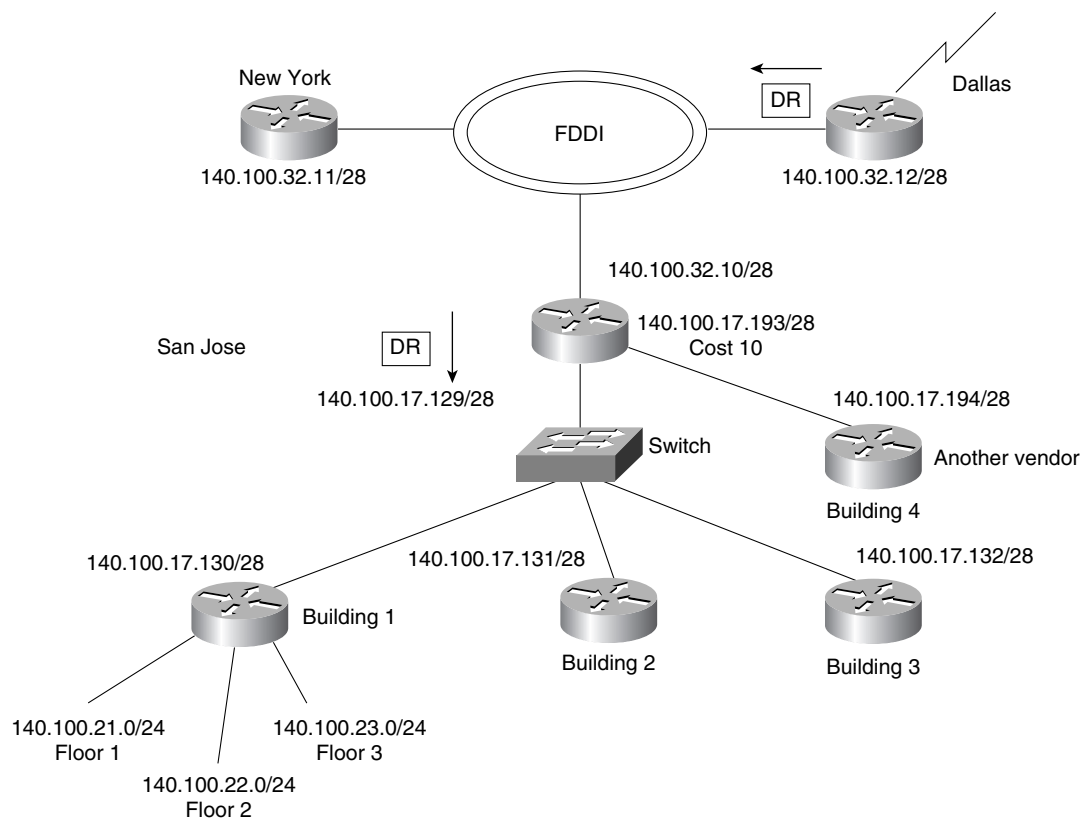


Table 7-3 explains the command.

Table 7-3 The `ip ospf network` Command

Option	Description
broadcast	Sets the network mode to broadcast.
non-broadcast	Sets the network mode to nonbroadcast multiaccess (NBMA mode). This is the default mode for serial interfaces with Frame Relay encapsulation and point-to-multipoint subinterfaces. Router(config-if)# ip ospf network non-broadcast
point-to-multipoint	Sets the network mode to point-to-multipoint.
point-to-multipoint non-broadcast	(Optional) Sets the network mode to point-to-multipoint nonbroadcast.

Configuring OSPF in NBMA Mode

In NBMA mode, the design considerations are imperative, because the selection of the DR and BDR must ensure physical connectivity to all routers in the NBMA cloud. Remember that this is a nonbroadcast environment, so you must configure the DR and BDR with a static list of the other routers attached to the cloud so that they can become neighbors and create adjacencies. This is achieved with the use of the **neighbor** command.

NOTE New technologies introduced to OSPF have removed the need to configure the neighbors manually. These technologies are explained later in the section “Configuring OSPF in Point-to-Multipoint Mode.”

The syntax of the command is as follows:

```
Router(config-if)#neighbor ip-address [priority number ] [poll-interval sec]
[ cost number]
```

You use the **neighbor** command to configure OSPF routers interconnecting to nonbroadcast networks. The different options used with the **neighbor** command are explained in Table 7-4.

Table 7-4 The OSPF neighbor Command

Syntax	Description
<i>ip-address</i>	Interface IP address of the neighbor.
priority number	(Optional) An 8-bit number indicating the likelihood of the neighbor being elected as the BDR or the DR. The default is 0. This keyword does not apply to point-to-multipoint mode interfaces because no BDR or DR is selected. This is another way of setting the ip ospf priority command. The highest priority is used, irrespective of the command used to set it.
poll-interval sec	(Optional) Unsigned integer value reflecting the poll interval. RFC 1247 recommends that this value be much larger than the hello interval. The default is 120 seconds (2 minutes). This keyword does not apply to point-to-multipoint mode interfaces. If a neighboring router has become inactive (hello packets have not been seen for the router dead interval period), it might still be necessary to send hello packets to the dead neighbor. These hello packets will be sent at a reduced rate, called the poll interval, to ensure connectivity is maintained while preserving bandwidth.
cost number	(Optional) Value that assigns a cost or metric. It takes the form of an integer from 1 to 65,535. Neighbors with no specific cost configured will assume the cost of the interface, based on the bandwidth or the ip ospf cost command.

Example 7-2 shows how the command is used.

Example 7-2 *The OSPF neighbor Command*

```
Router(config)#interface Serial0
Router(config-if)#ip address 131.144.10.100 255.255.255.0
Router(config-if)#encapsulation frame-relay
Router(config-if)#ip ospf network non-broadcast
Router(config)#router ospf 1
Router(config-router)#network 131.144.10.100 0.0.0.255 area 0
Router(config-router)#neighbor 131.144.10.2
Router(config-router)#neighbor 131.144.10.3
Router(config-router)#neighbor 131.144.10.5
```

NOTE NBMA mode is used by default in a nonbroadcast multiaccess environment, so there is no need for the **ip ospf network non-broadcast** command. However, **neighbor** statements are necessary.

Configuring OSPF in Point-to-Multipoint Mode

An OSPF point-to-multipoint interface is seen as a numbered point-to-point interface with one or more neighbors. The cloud is configured as one subnet with a host route for each router involved in the OSPF cloud.

NOTE The **ip ospf network point-to-multipoint non-broadcast** command option is a feature related to point-to-multipoint networks with Cisco IOS Software Release 11.3a. You can find more information on the subject by searching Cisco.com with the keywords “OSPF point-to-multipoint network with separate costs per neighbor” or click the link to the Technical Support OSPF page.

By default, the network is considered to be a series of point-to-point interfaces. There is no need to specify neighbors, because the neighbors will see each other and simply become adjacent, with no need for the election of a DR or a BDR. Point-to-multipoint does not try to reduce adjacencies using a DR. Instead, it accepts the extra overhead of having a full set of adjacencies for the sake of stability. Point-to-multipoint forms an adjacency automatically along any PVC, which causes more overhead but is more resilient than NBMA.

You can specify neighbors with the **neighbor** command, in which case you should specify a cost to each neighbor. You are not required to have a fully meshed topology, which reduces the number of PVCs needed and the number of neighbor entries in the neighbor table.

It is possible to change the default of NBMA to point-to-multipoint with the command **ip ospf network point-to-multipoint** and to nonbroadcast network with the command **ip ospf network point-to-multipoint non-broadcast**. The point-to-multipoint network is then considered a nonbroadcast network, and the mode is a Cisco extension. The **neighbor** command is required to identify neighbors in a nonbroadcast network. Assigning a cost to a neighbor is optional.

Example 7-3 shows the necessary configuration for OSPF in point-to-multipoint mode (point-to-multipoint broadcast mode, compliant with the RFC 2328, because the keyword **non-broadcast** is not specified). There is no need to configure neighbors, although you can do so if desired.

Example 7-3 *Configuring Point-to-Multipoint Networks*

```
Router(config)#interface Serial0
Router(config-if)#ip address 10.1.1.1 255.255.255.0
Router(config-if)#encapsulation frame-relay
Router(config-if)#ip ospf network point-to-multipoint
Router(config)#router ospf 1
Router(config-router)#network 10.1.1.0 0.0.0.255 area 0
```

Configuring OSPF in Broadcast Mode

You use the broadcast mode to avoid using the **neighbor** command and all the attendant configurations.

This broadcast mode works best with a fully meshed network. Example 7-4 shows a typical configuration of OSPF in broadcast mode.

Example 7-4 *Configuring a Broadcast Network*

```
Router(config)#interface Serial0
Router(config-if)#ip address 10.1.1.1 255.255.255.0
Router(config-if)#encapsulation frame-relay
Router(config-if)#ip ospf network broadcast
Router(config)#router ospf 1
Router(config-router)#network 10.1.1.0 0.0.0.255 area 0
```

Configuring OSPF in Point-to-Point Mode on a Frame Relay Subinterface

In the point-to-point mode, the adjacency created between the routers is automatic because each subinterface behaves as a physical point-to-point network. Therefore, the communication is direct and automatic.

The following steps explain how to configure OSPF point-to-point mode on subinterfaces:

Step 1 Configure Frame Relay encapsulation on the interface.

Step 2 At the interface level, create a subinterface.

It is recommended that you remove any network layer address assigned to the physical interface using the **no ip address** command. The Layer 3 address should be assigned to the subinterface.

Step 3 Configure the Layer 3 and Layer 2 (DLCI) addresses on the subinterface.

Step 4 Point-to-point mode is the default OSPF mode for point-to-point subinterfaces, so no further configuration is required.

Example 7-5 shows the necessary configuration required for a point-to-point Frame Relay subinterface.

Example 7-5 *Configuring a Point-to-Point Frame Relay Subinterface*

```
Router(config)#interface Serial0
Router(config-if)#no ip address
Router(config-if)#encapsulation frame-relay
Router(config)#interface Serial0.1 point-to-point
Router(config-subif)#ip address 10.1.1.1 255.255.255.0
Router(config-subif)#frame-relay interface-dlci 51
Router(config)#interface Serial0.2 point-to-point
Router(config-subif)#ip address 10.1.2.1 255.255.255.0
Router(config-subif)#frame-relay interface-dlci 52
Router(config)#router ospf 1
Router(config-router)#network 10.1.0.0 0.0.255.255 area 0
```

The shading in the previous example shows the configuration required to create a subinterface running point-to-point Frame Relay with IP.

Of course, it is imperative to check any configuration on a network device, because any errors could potentially bring down the entire network. To verify the configuration, there is a wealth of Cisco commands. They are covered in the following section.

Checking the Configuration of OSPF on a Single Router

The set of commands shown in Table 7-5 is invaluable in both configuration and maintenance of a live network. These commands are particularly useful in troubleshooting the network. As such, these commands are a necessary set of tools for use on a daily basis, for the CCNP/CCDP/CCIP BSCI exam, and for the CCIE lab exam.

Table 7-5 The **show** Command Options for OSPF

Command Option	Description
show ip ospf	Shows the OSPF process and its details, for example, how many times the router has recalculated its routing table.
show ip ospf database	Shows the contents of the topological database.
show ip ospf interface	Gives information about how OSPF has been configured on each interface. Typing errors are easily seen with this command.
show ip ospf neighbor	Displays all the information about the relationship that the router has with its neighbors—for example, the status of communication. One of the main points of interest is whether all the neighbors are present in the table.
show ip protocols	Enables you to view the IP routing protocol configuration on the router.
show ip route	Shows detailed information about the networks that the router is aware of and the preferred paths to those networks. Also gives the next logical hop as the next step in the path.

NOTE An excellent resource is the Cisco web site, which takes you through a series of flowcharts to identify a particular problem and how to solve it. These flowcharts are nested, so expect to work through several levels. This is an excellent resource for understanding troubleshooting techniques that are beyond the scope of this book or the BSCI exam. You can find the flowcharts at this URL:

http://www.cisco.com/warp/public/104/trouble_main.html

Understanding the output of these commands is important. This is not just because the output might constitute questions on the exam, but because the capability to analyze what is happening on the network demands a thorough understanding of the concepts explained in this chapter. You need to understand the concepts in this chapter to interpret the output of a **show** command.

The OSPF **show** commands are highly detailed and give a comprehensive understanding of the state of the network.

The show ip ospf Command

This section explains the **show ip ospf** command. This command is extremely useful, because it shows how the OSPF routing protocol is running on a particular router. It includes the number of times that the SPF routing algorithm has been run, which is indicative of the stability of the network. To issue the command, use the following syntax:

```
Router#show ip ospf [process-id]
```

Example 7-6 shows the output of this command. Table 7-6 explains how to read this information.

Example 7-6 *The show ip ospf process-id Command Output*

```
SanJose#show ip ospf 100
Routing Process "ospf 100" with ID 140.100.32.10
Supports only single TOS(TOS0) routes
It is an internal router
SPF schedule delay 5 secs, Hold time between two SPFs 10 secs
Minimum LSA interval 5 secs. Minimum LSA arrival 1 secs
Number of external LSA 0. Checksum Sum 0x0
Number of DCbitless external LSA 0
Number of DoNotAge external LSA 0
Number of areas in this router is 1. 1 normal 0 stub 0 nssa
  Area 3
    Number of interfaces in this area is 3
    Area has no authentication
    SPF algorithm executed 10 times
    Area ranges are
    Link State Update Interval is 00:30:00 and due in 00:18:54
    Link State Age Interval is 00:20:00 and due in 00:08:53
    Number of DCbitless LSA 2
    Number of indication LSA 0
    Number of DoNotAge LSA 0
```

Table 7-6 *Explanation of the show ip ospf Command Output*

Field	Explanation
Routing Process “ospf 100” with ID 140.100.32.10	Shows the local process ID for OSPF and the RID that it will advertise.
Supports only single ToS (TOS0) routes	OSPF is capable of carrying information about the type of service (ToS) that the IP datagram has requested. This is supported by Cisco in accordance with the RFCs but only implements the value 0.
It is an internal router	Species the types of router that OSPF defines, including internal, area border, and autonomous system boundary router.
SPF schedule delay	Specifies how long to wait to start the SPF calculation after receiving an LSA update to prevent running SPF too often.
Hold time between two SPFs	Specifies the minimum amount of time between SPF calculations.
Number of DCbitless external LSA	Used with OSPF demand circuits. Refer to the Cisco web site for greater detail on this subject.

(continues)

Table 7-6 Explanation of the **show ip ospf** Command Output (Continued)

Field	Explanation
Number of DoNotAge external LSA	Used with OSPF demand circuits, such as ISDN.
Summary Link update interval is 0:00:00 and the update is due in 0:00:00	An ABR would transmit this link-state advertisement (LSA) into another area. Summarization occurs at the area border. As an internal router, this router is not capable of issuing this update.
External Link update interval is 0:00:00 and the update due in 0:00:00	An autonomous system boundary router would transmit this LSA into another routing protocol using redistribution. The update is external to the domain or autonomous system. As an internal router, this router is not capable of issuing this update.
Area 3 Number of interfaces in this area is 3 Area has no authentication SPF algorithm executed 10 times Area ranges are	Specifies the number of areas of which this router is a member. As an internal router, it is configured for a single area and is a member of one area. At a glance, it is possible to see how many of the router's interfaces are in an area and whether the router is using MD5 security. It is useful to see the number of times that the SPF algorithm has been executed, because this is an indication of the network stability. The area ranges show any summarization that has been configured.
Link State Update Interval is 00:30:00 and due in 00:18:54	The default for the LSA update timer is 30 minutes. This is used to ensure the integrity of the topological databases. This field shows when the next update is and that the default has not been changed. These update timers should be the same throughout the area.
Link State Age Interval is 00:20:00 and due in 00:08:53	This specifies the MAX-AGED update deletion interval and shows when the database will next be purged of out-of-date routes.

The show ip ospf database Command

The following command displays the contents of the router's topological database and the different LSAs that have populated the database:

```
Router#show ip ospf database
```

In this example, because the router used is an internal router, the LSAs displayed will be the router and network updates. This command has many parameters that enable the user to examine very specific information. This section considers the general command.

Example 7-7 shows the output of this command. Table 7-7 explains the meaning of the important fields.

Example 7-7 *The show ip ospf database Command Output*

SanJose#show ip ospf database					
OSPF Router with ID (140.100.32.10) (Process ID 100)					
Router Link States (Area 3)					
Link ID	ADV Router	Age	Seq#	Checksum	Link count
140.100.17.131	140.100.17.131	471	0x80000008	0xA469	1
140.100.17.132	140.100.17.132	215	0x80000007	0xA467	1
140.100.17.194	140.100.17.194	1489	0x8000000B	0xFF16	1
140.100.23.1	140.100.23.1	505	0x80000006	0x56B3	1
140.100.32.10	140.100.32.10	512	0x8000000C	0x46BA	3
140.100.32.11	140.100.32.11	150	0x80000006	0x6A73	1
140.100.32.12	140.100.32.12	1135	0x80000002	0x8E30	1
Net Link States (Area 3)					
Link ID	ADV Router	Age	Seq#	Checksum	
140.100.17.130	140.100.23.1	220	0x80000007	0x3B42	
140.100.17.194	140.100.17.194	1490	0x80000002	0x15C9	
140.100.32.11	140.100.32.11	150	0x80000004	0x379E	

Table 7-7 *Explanation of the show ip ospf database Command*

Field	Explanation
OSPF Router with ID (140.100.32.10) (Process ID 100)	The RID and the process ID of the router being viewed.
Router Link States (Area 3)	The router LSAs, showing the links connecting the router to neighbors discovered via the Hello protocol.
Link ID	The link ID, which is the same as the OSPF RID.
ADV Router	The OSPF RID of the advertising router. Note that the ID is the same as the link ID when describing the router LSAs. This is because the router is advertising these links in its router LSA to the area.
Age	The age is the length of time since the last update. It is shown in seconds.

(continues)

Table 7-7 Explanation of the **show ip ospf database** Command (Continued)

Field	Explanation
Seq #	The sequence number, used to ensure that the LSA is truly an update that is more recent than anything currently in the topological database.
Checksum	The checksum on the entire LSA update. Ensures the integrity of the update.
Link count	The number of links that the router has configured for OSPF. Note that this field is shown only for the router LSA update.
Net Link States (Area 3)	Information taken from the network LSAs that have been received by the router.
Summary Net Link States (Area 3)	Information taken from the summary LSAs, which are passed between the ABRs. As an internal router in a single area, this section of the display would be blank.

The **show ip ospf interface** Command

This command shows how OSPF has been configured on an interface level and how it is working at the interface. This level of detail is excellent for troubleshooting configuration errors:

```
Router#show ip ospf interface [type number]
```

Important information such as the DR, the BDR, a list of neighbors, and the network type is shown by this command. Example 7-8 shows the output of this command. Table 7-8 explains how to read this information.

Example 7-8 The **show ip ospf interface** [type number] Command Output

```
SanJose#show ip ospf interface fastethernet1/0
FastEthernet1/0 is up, line protocol is up
  Internet Address 140.100.17.129/28, Area 3
  Process ID 100, Router ID 140.100.32.10, Network Type BROADCAST, Cost: 1
  Transmit Delay is 1 sec, State DR, Priority 100
  Designated Router (ID) 140.100.32.10, Interface address 140.100.17.129
  Backup Designated router (ID) 140.100.23.1, Interface address 140.100.17.130
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
    Hello due in 00:00:06
  Neighbor Count is 3, Adjacent neighbor count is 2
    Adjacent with neighbor 140.100.17.132
    Adjacent with neighbor 140.100.17.131
    Adjacent with neighbor 140.100.23.1 (Backup Designated Router)
  Suppress hello for 0 neighbor(s)
```

Table 7-8 *Explanation of the show ip ospf interface Command*

Field	Explanation
FastEthernet1/0 is up, line protocol is up	This should be seen as two statements. The first half of the sentence indicates that the physical line is operational. This meaning differs with the type of interface; for Ethernet, it indicates the presence of the transceiver. The second portion of the sentence indicates that the data link layer is working.
Internet Address 140.100.17.129/28,	The IP address and mask configured on the interface.
Area 3	The OSPF area for which the interface is configured.
Process ID 100, Router ID 140.100.32.10	The autonomous system number, which is in fact the OSPF process ID. The RID shown will be advertised in the LSA updates.
Network Type BROADCAST	The type of network to which the interface is connected, which indicates how neighbors are found and adjacencies are formed.
Cost: 1	The metric cost of the link, which, although not stated, was probably dynamically chosen using the Cisco defaults.
Transmit Delay is 1 sec	The anticipated time taken to send an update to the neighbor. The default is 1 second.
State DR	The state of the link in reference to establishing adjacencies. This field is extremely useful in troubleshooting. Here are the states in order of progression: DOWN—Heard from no one. ATTEMPT—Sent a hello on an NBMA, but haven't heard back. INIT—Heard a hello, but have not achieved neighbor status. TWO-WAY—Established full neighbor relationship; saw itself in the neighbor's hello table. EXSTART—Starting up the link for exchanging DDPs. EXCHANGE—Sending DDPs to other router. LOADING—Building the database and LSAs from the DDPs. FULL—Established adjacency. DR—Is the designated router for this LAN.

(continues)

Table 7-8 Explanation of the **show ip ospf interface** Command (Continued)

Field	Explanation
Priority 100	The priority is sent in the Hello protocol and is used to determine the election of the DR and the BDR. The value of 1 means that the router is prepared to be elected. If every other router has the priority of 1, the highest RID will select the routers.
Designated Router (ID) 140.100.32.10, Interface address 140.100.17.129	The address of the elected DR. Note that the ID and the interface ID differ. This is a useful field for troubleshooting misconfiguration.
Backup Designated router (ID) 140.100.23.1, Interface address 140.100.17.130	The address of the BDR. Note that both the ID and the interface are given, and that they differ.
Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5	It is possible to change these timers, and it is sometimes necessary if connecting to another vendor's equipment that has different defaults. These timers should be consistent throughout the area. The timer intervals shown here are the defaults for broadcast and point-to-point network types.
Hello due in 00:00:06	When the next hello packet is due to be sent out of the interface.
Neighbor Count is 3, Adjacent neighbor count is 2	The number of routers that have neighbor relationships. Note that the number of routers with which adjacency is established is less than the number of neighbors. This is because there is a DR and a BDR, whose responsibility it is to maintain the adjacencies with all routers on the LAN.
Adjacent with neighbor 140.100.23.1 (Backup Designated Router)	The RID of the adjacent router, which is the BDR in this case.

The show ip ospf neighbor Command

This command shows OSPF neighbors. All the neighbors known to the router can be viewed, or the command can be made more granular and the neighbors can be shown on a per-interface basis. One neighbor also might be picked out for scrutiny. This level of detail is excellent for troubleshooting configuration errors:

```
Router#show ip ospf neighbor [type number] [neighbor-id] [detail]
```


Example 7-9 shows the output of this command.

Example 7-9 *The show ip ospf neighbor Command Output*

SanJose#show ip ospf neighbor					
Neighbor ID	Pri	State	Dead Time	Address	Interface
140.100.17.132	1	FULL/DROTHER	00:00:36	140.100.17.132	FastEthernet1/0
140.100.17.131	1	FULL/DROTHER	00:00:37	140.100.17.131	FastEthernet1/0
140.100.23.1	1	FULL/BDR	00:00:38	140.100.17.130	FastEthernet1/0
140.100.32.12	1	FULL/DROTHER	00:00:35	140.100.32.12	Fddi2/0
140.100.32.11	1	FULL/DR	00:00:32	140.100.32.11	Fddi2/0
140.100.17.194	1	FULL/DR	00:00:31	140.100.17.194	FastEthernet3/0

To be more specific in what is viewed, it is possible to look at the neighbors that have been discovered on a particular interface, as seen in Example 7-10.

Example 7-10 *The Neighbors That Have Been Discovered on a Particular Interface*

SanJose#show ip ospf neighbor fddi 2/0					
Neighbor ID	Pri	State	Dead Time	Address	Interface
140.100.32.12	1	FULL/DROTHER	00:00:36	140.100.32.12	Fddi2/0
140.100.32.11	1	FULL/DR	00:00:32	140.100.32.11	Fddi2/0

To see all the neighbors in as much detail as possible, however, use the command displayed in Example 7-11.

Example 7-11 *Using the show ip ospf neighbor detail Command*

```
SanJose#show ip ospf neighbor detail
Neighbor 140.100.17.132, interface address 140.100.17.132
  In the area 3 via interface FastEthernet1/0
  Neighbor priority is 1, State is FULL, 6 state changes
  DR is 140.100.17.129 BDR is 140.100.17.130
  Options 2
  Dead timer due in 00:00:35
Neighbor 140.100.17.131, interface address 140.100.17.131
  In the area 3 via interface FastEthernet1/0
  Neighbor priority is 1, State is FULL, 6 state changes
  DR is 140.100.17.129 BDR is 140.100.17.130
  Options 2
  Dead timer due in 00:00:34
Neighbor 140.100.23.1, interface address 140.100.17.130
  In the area 3 via interface FastEthernet1/0
  Neighbor priority is 1, State is FULL, 6 state changes
  DR is 140.100.17.129 BDR is 140.100.17.130
```

continues

Example 7-11 Using the `show ip ospf neighbor detail` Command (Continued)

```
Options 2
Dead timer due in 00:00:36
Neighbor 140.100.32.12, interface address 140.100.32.12
In the area 3 via interface Fddi2/0
Neighbor priority is 1, State is FULL, 6 state changes
DR is 140.100.32.11 BDR is 140.100.32.10
Options 2
Dead timer due in 00:00:32
Neighbor 140.100.32.11, interface address 140.100.32.11
In the area 3 via interface Fddi2/0
Neighbor priority is 1, State is FULL, 6 state changes
DR is 140.100.32.11 BDR is 140.100.32.10
Options 2
Dead timer due in 00:00:38
Neighbor 140.100.17.194, interface address 140.100.17.194
In the area 3 via interface FastEthernet3/0
Neighbor priority is 1, State is FULL, 9 state changes
DR is 140.100.17.194 BDR is 140.100.17.193
Options 2
Dead timer due in 00:00:38
```

Table 7-9 explains the meanings of the important fields from Examples 7-9 through 7-11.

Table 7-9 Explanation of the `show ip ospf neighbor` Command

Field	Explanation
Neighbor	This is the RID.
Neighbor priority	This is the priority sent out with the Hello protocol to elect the DR and the BDR.
State	This shows the state, not of the link, but whether the interface was elected. DR—Designated router. BDR—Backup designated router. DROTHER—The router was not chosen as the DR or the BDR. If the priority on the interface had been set to zero, the state would always be DROTHER because the router could not be elected as a DR or a BDR.
Dead Time	The dead time is how long the router will wait without hearing the periodic hello from its neighbor before it is declared dead. This timer should be consistent on the network; otherwise, there will be problems.

Table 7-9 *Explanation of the show ip ospf neighbor Command (Continued)*

Field	Explanation
Address	This is the interface address of the neighbor. Note that the RID is not the same as the interface address. If the loopback address or the highest IP address on the router has been used, the address probably will differ.
Interface	This is the outgoing interface of the router, upon which the neighbor routers were heard.
Options	The option available is one of design. It identifies whether the area the neighbors inhabit is a stub area.

The show ip protocols Command

This command shows the configuration of IP routing protocols on the router. It details how the protocols were configured and how they interact with one another. It also indicates when the next updates will occur. This command is excellent for troubleshooting configuration errors and understanding how the network is communicating about its routes:

```
Router#show ip protocols
```

Example 7-12 shows the output of this command. Table 7-10 explains how to read this information.

Example 7-12 *The show ip protocols Command Output*

```
SanJose#show ip protocols
Routing Protocol is "ospf 100"
  Sending updates every 0 seconds
  Invalid after 0 seconds, hold down 0, flushed after 0
  Outgoing update filter list for all interfaces is not set
  Incoming update filter list for all interfaces is not set
  Redistributing: ospf 100
  Routing for Networks:
    140.100.0.0
  Routing Information Sources:
    Gateway         Distance      Last Update
    140.100.17.131   110          00:50:23
    140.100.17.132   110          00:50:23
    140.100.17.194   110          00:07:39
    140.100.23.1     110          00:50:23
    140.100.32.11    110          00:07:39
    140.100.32.12    110          00:07:39
  Distance: (default is 110)
```

Table 7-10 *Explanation of the show ip protocols Command*

Field	Explanation
Routing Protocol is “ospf 100”	This routing protocol is configured on the router. If there is more than one routing protocol configured, the details of each are listed in turn.
Sending updates every 0 seconds	The frequency of the routing update is shown. It is not relevant for a link-state routing protocol that sends updates of changes as required (incremental updates).
Invalid after 0 seconds	This field is relevant for distance vector protocols. It indicates the period of time that a route is considered valid, from the time of the last update. If an update on the status of the route has not been received in this defined value, the route is marked unreachable.
hold down 0	Holddown timers are used only in distance vector protocols. If a distance vector protocol suspects that a route in its table is bad, it will mark it down but will not accept another path with a less-favorable metric until the holddown timer has expired. This is to avoid loops in the network. If a link-state protocol hears an update, it acts on the information.
flushed after 0	The 0 value indicates that this is a field used by distance vector routing protocols. After marking a route as invalid, it will flush it from the routing table after this timer has expired.
Outgoing update filter list for all interfaces is not set	Access lists can be set on an interface to filter networks from the routing update. This should be used carefully because it affects connectivity.
Incoming update filter list for all interfaces is not set	The access list can filter either outgoing or incoming updates.
Redistributing: ospf 100	If the routing protocol is sharing information with another routing protocol configured on the router, the information is listed here. This is a very important field because redistribution is complex and, therefore, easily misconfigured. If no redistribution is configured, the protocol is seen to be sharing information with itself.
Routing for Networks: 140.100.0.0	This reflects the use of the network commands when the protocol was configured. OSPF allows granularity in the use of the command. The entries here could be as specific as the interface addresses.
Routing Information Sources	This is a major heading for the gateway fields, which are the addresses of the routers sending updates to this router. They will become the next logical hop in the routing table.

Table 7-10 *Explanation of the **show ip protocols** Command (Continued)*

Field	Explanation
Gateway	This field is a subset of the Routing Information Sources field just discussed. It is the address of the router providing updates.
Distance	The administrative distance is the preference given to the source of the update. Whereas the metric indicates which path to choose if there is more than one available, the administrative distance indicates which source (routing protocol) to choose if there is more than one providing a path to a remote network. The administrative distance takes precedence over the routing metric.
Last Update	This is the time since the last update was received from that source.
Distance: (default is 110)	The administrative distance can be changed for the entire routing protocol (the example here is OSPF), which would be listed here, or it can be changed per source, as seen earlier in the listing of each individual source (gateway).

The **show ip route** Command

This command shows the IP routing table on the router. It details how the network is known to the router and how the router discovered the route. This command is excellent for troubleshooting configuration errors and understanding how the network is communicating about its routes. It is given detailed consideration in Chapter 1, “IP Routing Principles.”

Router#**show ip route**

The commands covered in this section are useful to verify that the configuration has worked and that the OSPF network is functioning correctly. In a single-area environment, the full complexity of OSPF is not engaged. The full strength and complexity of OSPF come to the forefront in the design and configuration of a multiarea network.

Troubleshooting OSPF in a Single Area

Troubleshooting an OSPF network requires the same skills of detection and critical thinking as those needed in any problem solving. Whether you are a doctor trying to locate the cause of a patient’s pain or a network administrator investigating a corporate network’s slow response time, the approach needs to be methodical and well-documented. The better your understanding of the subject, or patient, the easier it is to diagnose the problem. As a network administrator, this understanding will grow if you document not only your network, but also every change that is made to that network.

Cisco provides many tools to aid the troubleshooting process. There are courses and many technical documents in addition to a forum for questions and answers. The following web pages provide some excellent troubleshooting tools, which will help in the everyday maintenance of your network and provide a good learning resource. Some of these resources might require you to be a registered user and to log in:

- A page of troubleshooting tools:
<http://www.cisco.com/public/support/tac/tools.shtml>

Because Cisco sometimes reorganizes its web pages, this reference might change. Luckily, Cisco maintains an excellent search engine. You should search for tools from the home page.
- A utility that allows the output of listed show commands and interprets the output:
<https://www.cisco.com/cgi-bin/Support/OutputInterpreter/home.pl>
- A troubleshooting assistant that leads you through a series of screens to help diagnose a problem:
http://www.cisco.com/kobayashi/support/tac/tsa/launch_tsa.html

The debug Commands

An excellent, though dangerous, troubleshooting tool is the **debug** command. The **debug** command has the highest process priority and is therefore capable of consuming all the resources on the router, thus becoming the problem as opposed to helping to solve the problem. It is important simply to turn on debug for the specific task to be monitored and to turn it off as soon as the data is gathered. The **no** form of this command disables debugging output.

You would be wise to direct the output to a log file so that the data can be perused with care. The **debug** commands are EXEC commands.

The options available for monitoring OSPF are listed in Table 7-11.

Table 7-11 The **debug** Command Options for OSPF

Command Option	Description
debug ip ospf events	Displays information about OSPF-related events, such as adjacencies, flooding information, designated router selection, and shortest path first (SPF) calculation.
debug ip packet	IP debugging information includes packets received, generated, and forwarded. Fast-switched packets do not generate messages.

Foundation Summary

The “Foundation Summary” section of each chapter lists the most important facts from the chapter. Although this section does not list every fact from the chapter that will be on your exam, a well-prepared candidate should, at a minimum, know all the details in each “Foundation Summary” before going to take the exam.

The following is a list of the commands explained in this chapter. This list is not intended to teach the use of the commands, but to remind you of the options available.

The commands in Table 7-12 are configuration commands.

Table 7-12 OSPF Configuration Command Summary

Command	Description
router ospf <i>process-number</i>	Turns on the OSPF process and identifies it with a process ID.
network <i>network-number wildcard-mask</i> area <i>area-number</i>	Identifies which networks—and, thus, interfaces—belong to which area.
interface loopback <i>interface-number</i>	Creates the loopback interface that can now be used to create the RID.
ip ospf cost <i>cost</i>	Sets the cost or metric for the outgoing interface.
auto-cost reference-bandwidth <i>reference-bandwidth</i>	Allows the administrator to change the formula used to calculate the metric.
ip ospf priority <i>number</i>	Helps to determine which router on a multiaccess network will be elected as the DR.
ip ospf network <i>network-type</i>	Identifies the type of network to the OSPF process, which will determine how the adjacencies are created.
neighbor <i>ip-address</i> [priority <i>number</i>] [poll-interval <i>sec</i>] [cost <i>number</i>]	States the neighbor, its address, and its capability to become the DR. This command also allows the configuration of cost and the timers on the interface.
router-id <i>ip-address</i>	Used to define the RID of the router.

Table 7-13 The *show* Command Options for OSPF

Command Option	Description
show ip ospf	Shows the OSPF process and its details, for example, how many times the router has recalculated its routing table.
show ip ospf database	Shows the contents of the topological database.
show ip ospf interface	Gives information about how OSPF has been configured on each interface. Typing errors are easily seen with this command.
show ip ospf neighbor	Displays all the information about the relationship that the router has with its neighbors—for example, the status of communication. One of the main points of interest is whether all the neighbors are present in the table.
show ip protocols	Enables you to view the IP routing protocol configuration on the router.
show ip route	Shows detailed information about the networks of which the router is aware and the preferred paths to those networks. Also gives the next logical hop as the next step in the path.

The **debug** commands are EXEC commands. The options available for monitoring OSPF are listed in Table 7-14.

Table 7-14 The *debug* Command Options for OSPF

Command Option	Description
Debug ip ospf events	Displays information about OSPF-related events, such as adjacencies, flooding information, DR selection, and shortest path first (SPF) calculation.
debug ip packet	IP debugging information includes packets received, generated, and forwarded. Fast-switched packets do not generate messages.

Q&A

As mentioned in the introduction, “All About the CCNP, CCDP, and CCIP Certifications,” you have two choices for review questions. The questions that follow next give you a bigger challenge than the exam itself by using an open-ended question format. By reviewing now with this more difficult question format, you can exercise your memory better and prove your conceptual and factual knowledge of this chapter. The answers to these questions are found in Appendix A.

For more practice with examlike question formats, including questions using a router simulator and multichoice questions, use the exam engine on the CD.

1. What command is used to manually determine which router on a LAN will become the DR?
2. What parameter is used to calculate the default metric of a route in OSPF on a Cisco router?
3. It is possible to have more than one OSPF process on a router. How would you do this?
4. Explain the command **ip ospf network non-broadcast** .
5. In which of the NBMA configuration choices is it necessary to state the neighbors manually? Why is this necessary?
6. In a Frame Relay environment, which is fully meshed, which OSPF configurations might be chosen? Give reasons for your choice.
7. The Cisco solution point-to-point mode does not require the configuration of DR and BDR. Explain briefly why.
8. The address 192.100.56.10/21 has been allocated to an interface on the router. This interface alone is to be included in the OSPF process. State the command that would start the process on this interface.
9. The metric used by OSPF is cost. How would you change the metric on an interface?
10. If the command **ip ospf network non-broadcast** is used, what additional statement is necessary?
11. What command shows which router on a LAN is the BDR?
12. Explain briefly what **show ip ospf database** will reveal.
13. What command is used to show the state of adjacencies?
14. Which command is used to show OSPF packets being sent and received in real time?
15. How would you show the OSPF process ID of the router?

16. What makes **debug** a dangerous command for your router?
17. What is the sequence number and where is it held?
18. In the **show ip ospf** command, there is a field called the SPF schedule delay. What is the purpose of this field, and what is the default time?
19. What is the advantage of the command **show ip ospf interface** in troubleshooting?

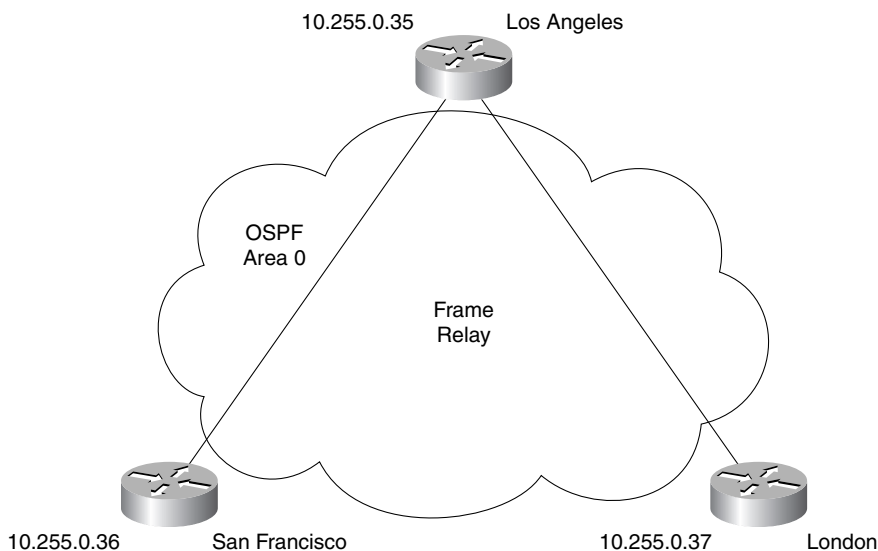
Scenarios

The following scenario and questions are designed to draw together the content of the chapter and to exercise your understanding of the concepts. There is not necessarily a right answer. The thought process and practice in manipulating the concepts is the goal of this section. The answers to the scenario questions are found at the end of this chapter.

Scenario 7-1

The company Jackanory.com is charged with configuring the Los Angeles site for OSPF. Figure 7-2 is a network diagram for this network. You must first understand the design requirements for the entire network. As the figure shows, Jackanory.com intends to use Frame Relay to connect the different sites. The company also has decided to use the private address of 10.0.0.0. The Los Angeles site is to be allocated 10.1.0.0 255.255.0.0, which can be subdivided. Referring to Figure 7-2, answer the following questions.

Figure 7-2 Network Diagram for Scenario 7-1



1. Design the addressing scheme for the site so that summarization can be implemented at a later stage.
2. Determine the configuration that will be implemented across the Frame Relay cloud.
3. State the configuration commands required to implement OSPF on the routers. Note that all the routers are to be in Area 0.

Scenario Answers

The answers provided in this section are not necessarily the only possible answers to the questions. The questions are designed to test your knowledge and to give practical exercise in certain key areas. This section is intended to test and exercise skills and concepts detailed in the body of this chapter.

If your answer is different, ask yourself whether it follows the tenets explained in the answers provided. Your answer is correct not if it matches the solution provided in the book, but rather if it has included the principles of design laid out in the chapter.

In this way, the testing provided in these scenarios is deeper: It examines not only your knowledge, but also your understanding and ability to apply that knowledge to problems.

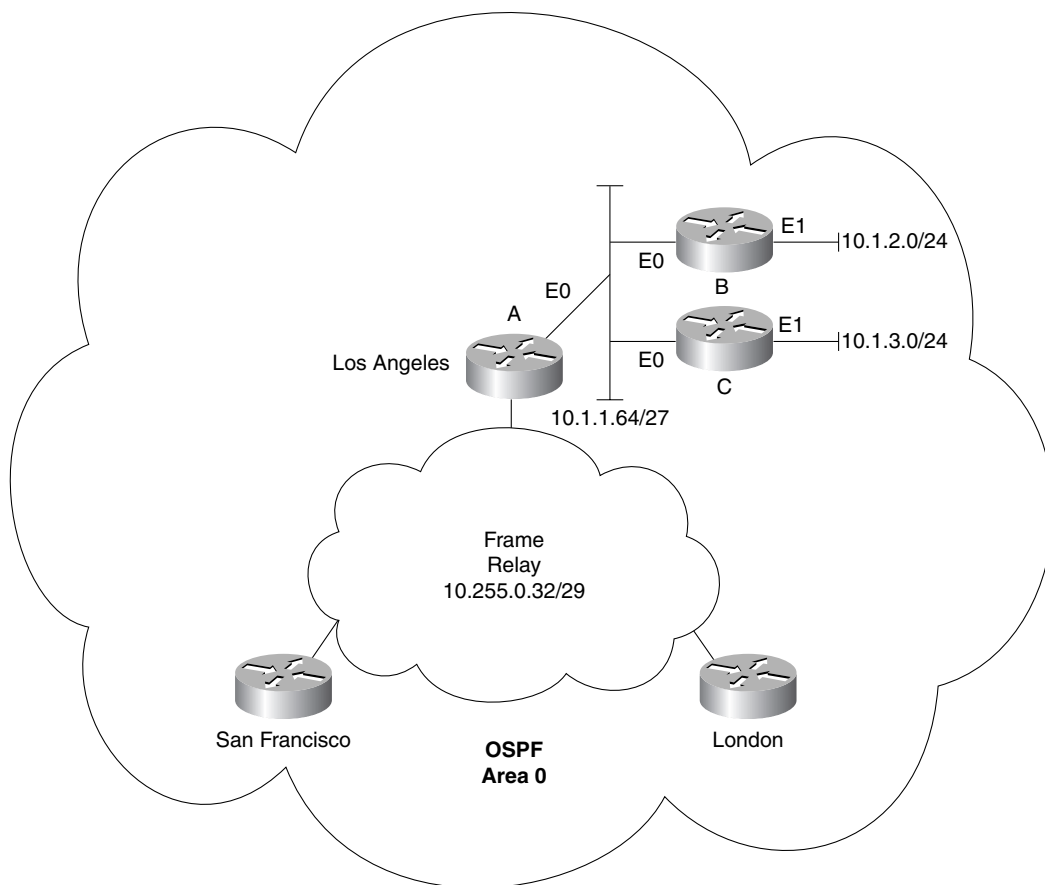
If you do not get the correct answer, refer back to the text and review the subject tested. Be certain to also review your notes on the question to ensure that you understand the principles of the subject.

Scenario 7-1 Answers

1. *Design the addressing scheme for the site so that summarization can be implemented at a later stage.*

An addressing scheme for the Los Angeles site could be as shown in Figure 7-3.

Figure 7-3 Scenario 7-1 Answer



2. *Determine the configuration that will be implemented across the Frame Relay cloud.*

The Frame Relay configuration is a hub-and-spoke topology. Because Los Angeles is the central hub of the company, it is logical that it is also the network hub.

The configuration in Frame Relay would therefore be a point-to-multipoint one. This avoids the DR/BDR election process, because it treats all the PVC links as a collection of point-to-point links, forming adjacencies across each link even though they are not separate point-to-point subinterfaces. Having separate links removes the need for the DR/BDR. This is a good solution if three different companies have a multivendor environment, because it conforms to the RFC standards.

Example 7-13 shows a configuration that could be implemented.

Example 7-13 *Answer to Scenario 7-1 Configuration Question 2*

```
Los Angeles
interface serial 0
encapsulation frame-relay
ip address 10.255.0.35 255.255.255.248
ip ospf network point-to-multipoint

router ospf 100
network 10.0.0.0 0.255.255.255 area 0

San Francisco
interface serial 0
encapsulation frame-relay
ip address 10.255.0.36 255.255.255.248
ip ospf network point-to-multipoint

router ospf 100
network 10.0.0.0 0.255.255.255 area 0

London
interface serial 0
encapsulation frame-relay
ip address 10.255.0.37 255.255.255.248
ip ospf network point-to-multipoint

router ospf 100
network 10.0.0.0 0.255.255.255 area 0
```

If individual point-to-point networks are chosen, then a separate PVC and IP subnet will be required for the configuration. This is illustrated in Example 7-14.

Example 7-14 *Answer to Scenario 7-1 Configuration Question 2*

```
Los Angeles
interface serial 0
no IP address
encapsulation frame-relay
interface serial0.1 point-to-point
ip address 10.255.0.35 255.255.255.252
frame-relay interface-dlci 21
interface serial0.1 point-to-point
ip address 10.255.0.49 255.255.255.252
frame-relay interface-dlci 28

router ospf 100
```

Example 7-14 *Answer to Scenario 7-1 Configuration Question 2 (Continued)*

```

network 10.0.0.0 0.255.255.255 area 0

San Francisco
interface serial 0
no IP address
encapsulation frame-relay
interface serial0.1 point-to-point
ip address 10.255.0.36 255.255.255.252
frame-relay interface-dlci 44

router ospf 100
network 10.0.0.0 0.255.255.255 area 0

London
interface serial 0
no IP address
encapsulation frame-relay
interface serial0.1 point-to-point
ip address 10.255.0.50 255.255.255.252
frame-relay interface-dlci 66

router ospf 100
network 10.0.0.0 0.255.255.255 area 0

```

Note that the Frame Relay cloud is a continuation of Area 0. This is a logical first step; it is advisable to change this configuration when OSPF is being run throughout the organization. When the other sites are up and running OSPF, the Los Angeles site could become another area. This would allow the summarization of routes across the Frame Relay cloud, which would reduce the traffic, cost, and possibility for congestion.

3. *State the configuration commands required to implement OSPF on the routers. Note that all the routers are to be in Area 0.*

The configuration on the Los Angeles routers, not including the Frame Relay configuration, is shown in Example 7-15.

Example 7-15 *Answer to Scenario 7-1 Configuration Question 3*

```

Router A:
router ospf 100
network 10.0.0.0 area 0

interface Ethernet 0
ip address 10.1.1.65 255.255.255.224
ip ospf priority 10

```

continues

Example 7-15 *Answer to Scenario 7-1 Configuration Question 3 (Continued)*

```
Router B:
router ospf 100
network 10.0.0.0 0.255.255.25 area 0

interface Ethernet 0
ip address 10.1.1.66 255.255.255.224

interface Ethernet 1
ip address 10.1.2.1 255.255.255.0

Router C
router ospf 100
network 10.0.0.0 0.255.255.255 area 0

interface ethernet 0
ip address 10.1.1.67 255.255.255.224

interface ethernet 1
ip address 10.1.3.0 255.255.255.0
```

Note that Router A has been given a priority of 10, ensuring that it becomes the DR. Although this is not essential, it is a clear configuration and is advisable because this router is a larger system and is very reliable. For the same reasons, it is also considered good system management to define the priority for the BDR.



This chapter covers the following topics, which you need to understand to pass the CCNP/CCDP/CCIP BSCI exam:

- The purpose of using OSPF in a multiple area network
- The features of multiple area OSPF
- The operation of OSPF across multiple areas
- Design considerations in multiple area OSPF

Using OSPF Across Multiple Areas

The topics in this chapter detail the routing protocol OSPF across multiple areas. This chapter assumes your knowledge of Chapter 6, “Using OSPF in a Single Area,” and Chapter 7, “Configuring OSPF in a Single Area,” which dealt with OSPF concepts and configuration in a single area. Chapter 8 builds on this understanding and explains how OSPF works within a large multiarea network. This chapter and the following chapter cover two major sections. Chapter 8 deals with how the protocol works theoretically. Chapter 9, “Configuring OSPF Across Multiple Areas,” covers how to implement and manage an OSPF network. This chapter introduces OSPF areas and explains the operation of the protocol across those areas.

Before you can configure OSPF in multiple areas, you need to understand the motivation for using OSPF in multiple areas. You then must understand how to determine the area boundaries in OSPF. In order to design a multiarea OSPF network in this way, you need a comprehensive grasp of the features of multiarea OSPF, the operation of OSPF across multiple areas and, of course, the design considerations of such a network. This chapter discusses each of these topics.

“Do I Know This Already?” Quiz

The purpose of the “Do I Know This Already?” quiz is to help you decide what parts of this chapter to use. If you already intend to read the entire chapter, you do not necessarily need to answer these questions now.

The 12-question quiz, derived from the major sections in “Foundation Topics” portion of the chapter, helps you to determine how to spend your limited study time.

Table 8-1 outlines the major topics discussed in this chapter and the “Do I Know This Already?” quiz questions that correspond to those topics.

Table 8-1 “Do I Know This Already?” Foundation Topics Section-to-Question Mapping

Foundation Topics Section	Questions Covered in This Section
The Purpose of Using OSPF in a Multiple Area Network	1–3
The Features of Multiple Area OSPF	4–6
The Operation of OSPF Across Multiple Areas	7–9
Design Considerations in Multiple Area OSPF	10–12

CAUTION The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark this question wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

- When using OSPF, which of the following is a problem with large routing tables?
 - The routing table is sent out every 30 seconds, which will cause congestion.
 - As the routing table increases in size, the longer each lookup becomes. The memory requirements on the router also increase.
 - The process will time out before a path to the destination is found.
 - Fast switching cannot be used if the routing table exceeds a certain size.
- Why does the frequency of the SPF algorithm increase with the size of the area?
 - Each router will have more neighbors, and synchronizing the databases takes priority.
 - The topology databases will subdivide after a certain size, requiring multiple SPF calculations.
 - As the area size increases, each recalculation will also take longer, which might result in errors, requiring the algorithm to rerun.
 - The larger the network, the greater the probability of a network change and, thus, a recalculation of the entire area.

3. Multiple areas are one of the main features of OSPF. Which of the following statements explain why this feature is such an important enhancement on earlier routing protocols?
 - a. It is easier to implement security.
 - b. All computation is kept within the area, with minimum communication between the areas, allowing the network to scale to larger sizes.
 - c. The network domain, when divided into areas allows for the use of both IANA classful addressing and private addressing.
 - d. The use of multiple areas allows for the use of prioritization.
4. What is an internal router?
 - a. An OSPF process running on a multilayer switch.
 - b. When multiple OSPF processes are run on the same router, the internal router is responsible for managing the processes.
 - c. A router responsible for maintaining a current and accurate database of every subnet within the area. All interfaces on this router are within the same area.
 - d. A router running OSPF with no external links to another autonomous system.
5. What is the purpose of an ABR?
 - a. A router responsible for connecting to outside the autonomous system
 - b. A router responsible for connecting two or more areas
 - c. A logical group of nodes forming a backbone that connects other areas
 - d. A group of routers running OSPF with no external links
6. What do the initials ASBR represent?
 - a. Authority Subnet Boundary Router
 - b. Autonomous System Border Router
 - c. Automatic Summarization Boundary Router
 - d. Autonomous System Boundary Router
7. How are routes that are generated within an area propagated throughout the area?
 - a. Type 3 and 4 LSAs
 - b. In the summary LSA sent out every 30 minutes
 - c. In the Hello packet between neighbors
 - d. Type 1 and 2 LSAs

8. Which of the following conditions must be met before any LSAs can be flooded out of all the interfaces?
 - a. The interface is in a state of exchange or full adjacency.
 - b. The interface is not connected to a totally stubby area.
 - c. The LSA was not received through the interface.
 - d. The interface is connected to a totally stubby area.
9. Which of the following OSPF characteristics affect how the routing table is created?
 - a. Whether there are multiple areas in the domain
 - b. Whether MD-5 security has been configured
 - c. The type of area in which the router is located
 - d. Whether there are communications outside the autonomous system
10. Which is the best design for OSPF?
 - a. Hierarchical with summarization
 - b. Tiered
 - c. Flat with summarization
 - d. Elliptical
11. Why does the type of area determine the number of routers that can be placed in the area?
 - a. Each LSA packet type has a fixed header, limiting the number of paths that can be listed.
 - b. The area type determines the number of LSAs and how often and how much CPU and memory each SPF computation requires.
 - c. Stub areas are not summarized, requiring additional resources.
 - d. The backbone area requires fewer resources because it simply sends summarized path information into other areas, which does not require the SPF algorithm to be run.
12. Which of the following must be observed when creating a virtual link?
 - a. Both routers must share a common area.
 - b. Both routers must share the same subnet address.
 - c. One of the routers must be connected to area 0.
 - d. Both routers must share the same process ID.

The answers to this quiz are found in Appendix A, “Answers to Chapter ‘Do I Know This Already?’ Quizzes and Q&A Sections.” The suggested choices for your next step are as follows:

- **6 or less overall score**—Read the entire chapter. This includes the “Foundation Topics” and “Foundation Summary” sections, the “Q&A” section, and the “Scenarios” at the end of the chapter.
- **7–9 overall score** —Begin with the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. If you have trouble with these exercises, read the appropriate sections in “Foundation Topics.”
- **10 or more overall score** —If you want more review on these topics, skip to the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. Otherwise, move to the next chapter.

Foundation Topics

The Purpose of Using OSPF in a Multiple Area Network

This section explains what multiple area networks are and how they overcome some of the shortcomings of single area networks. Multiple areas in OSPF provide one of the main distinguishing features between the distance vector protocols and the link-state OSPF.

As you learned in Chapter 6, an OSPF area is a logical grouping of routers that are running OSPF with identical topological databases. An area is a subdivision of the greater OSPF domain, sometimes known as the autonomous system. Multiple areas prevent a large network from outgrowing its capacity to communicate the details of the network to the routing devices charged with maintaining control and connectivity throughout the network.

The division of the autonomous system into areas allows routers in each area to maintain their own topological databases. This limits the size of the topological databases, and summary and external links ensure connectivity between areas and networks outside the autonomous system.

Problems with OSPF in a Single Area

To understand the true benefits of multiple areas, consider why you might decide to create multiple areas from one area.

The following symptoms that you might observe on the network provide a clue that a single area is becoming overpowered:

- The SPF algorithm is running more frequently. The larger the network, the greater the probability of a network change and, thus, a recalculation of the entire area. Each recalculation also takes longer.
- The larger the area, the greater the size of the routing table. The routing table is *not* sent out wholesale, as in a distance vector routing protocol; however, the greater the size of the table, the longer each lookup becomes. The memory requirements on the router also increase.
- The topological database increases in size and eventually becomes unmanageable for the same reasons as in the previous point. The topology table is exchanged between adjacent routers at least every 30 minutes.

As the various databases increase in size and the calculations become increasingly frequent, the CPU utilization increases while the available memory decreases. This will make the network response time sluggish (not because of congestion on the line, but because of congestion within the

router itself). It can also cause congestion on the link. These can result in various additional problems, such as loss of connectivity, loss of packets, and system hangs.

NOTE To check the CPU utilization on the router, use the **show processes cpu** command. To check the memory utilization, issue the **show memory free** command.

How to Determine Area Boundaries

Although you might have an obvious need for multiple areas, the practical question is how you should implement multiple areas. There are two approaches, as follows:

- To grow a single area until it becomes unmanageable
- To design the network with multiple areas, which are very small, in the expectation that the networks will grow to fit comfortably into their areas

Both approaches are valid. The first approach requires less initial work and configuration. Great care should be put into the design of the network, however, because this can cause problems in the future, particularly in addressing.

In practice, many companies convert their networks to OSPF from a distance vector routing protocol when they realize that they have outgrown the existing routing protocol. This allows the planned implementation of the second approach.

The Features of Multiple Area OSPF

Now that you understand why you need to control the size of the areas, you should consider the design issues for the different areas, including the technology that underpins them and their communication (both within and between the areas).

OSPF Within an Area

One of the main strengths of OSPF is its capability to scale and to support large networks. It does so by creating areas from groups of subnets. The area is seen internally almost as if it were a small organization or entity of its own. It communicates with the other areas, exchanging routing information; this exchange is kept to a minimum, however, allowing only that which is required for connectivity. All computation is kept within the area.

In this way, a router is not overwhelmed by the entirety of the organization's network. This is crucial, because the nature of a link-state routing protocol is more CPU- and memory-intensive.

Router Types

Given the hierarchical nature of the OSPF network, there are routers operating within an area, routers connecting areas, and routers connecting the organization or autonomous system to the outside world. Each of these routers has a different set of responsibilities, depending on its position and function within the OSPF hierarchical design.

The following list identifies the different OSPF routers:

- **Internal router** —Within an area, the functionality of the router is straightforward. It is responsible for maintaining a current and accurate database of every subnet within the area. It is also responsible for forwarding data to other networks by the shortest path. Flooding of routing updates is confined to the area. All interfaces on this router are within the same area. This router is the only router that can operate in a single area OSPF network, other than an Autonomous System Boundary Router (ASBR).
- **Backbone router** —The design rules for OSPF require that all the areas be connected through a single area, known as the *backbone area*, *Area 0*, or 0.0.0.0. A router within this area is referred to as a *backbone router*. It can also be an internal router, an ASBR, or an Area Border Router (ABR).
- **ABR**—This router is responsible for connecting two or more areas. It holds a full topological database for each area to which it is connected and sends LSA updates between the areas. These LSA updates are summary updates of the subnets within an area. Summarization should be configured for OSPF at the area border because this is where the LSAs make use of the reduced routing updates to minimize the routing overhead on both the network and the routers.
- **ASBR**—To connect to the outside world or to any other routing protocol, you need to leave the OSPF domain. OSPF is an *interior routing protocol* or *Interior Gateway Protocol* (IGP); *gateway* is an older term for a router. The router configured for this duty is the ASBR. If any routing protocols are being redistributed to OSPF on a router, the router will become an ASBR because the other routing protocols are outside the OSPF autonomous systems. Although you can place this router anywhere in the OSPF hierarchical design, it should reside in the backbone area. Because any traffic leaving the OSPF domain is also likely to leave the router's area, it makes sense to place the ASBR in a central location that all traffic leaving its area must traverse. This router could be configured within a single OSPF area, pointing to the outside world.

Figure 8-1 shows how the different router types are interrelated. All the routers in the backbone area, area 0, are not only performing the function of ABR, or ASBR as labeled, but are also backbone routers.

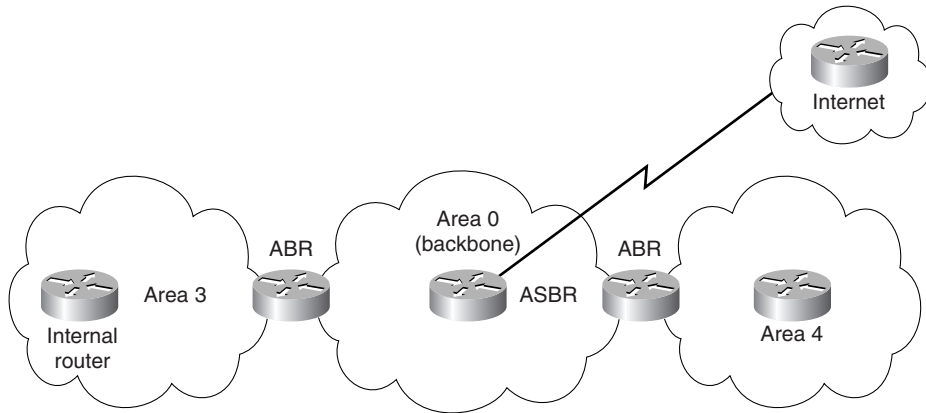
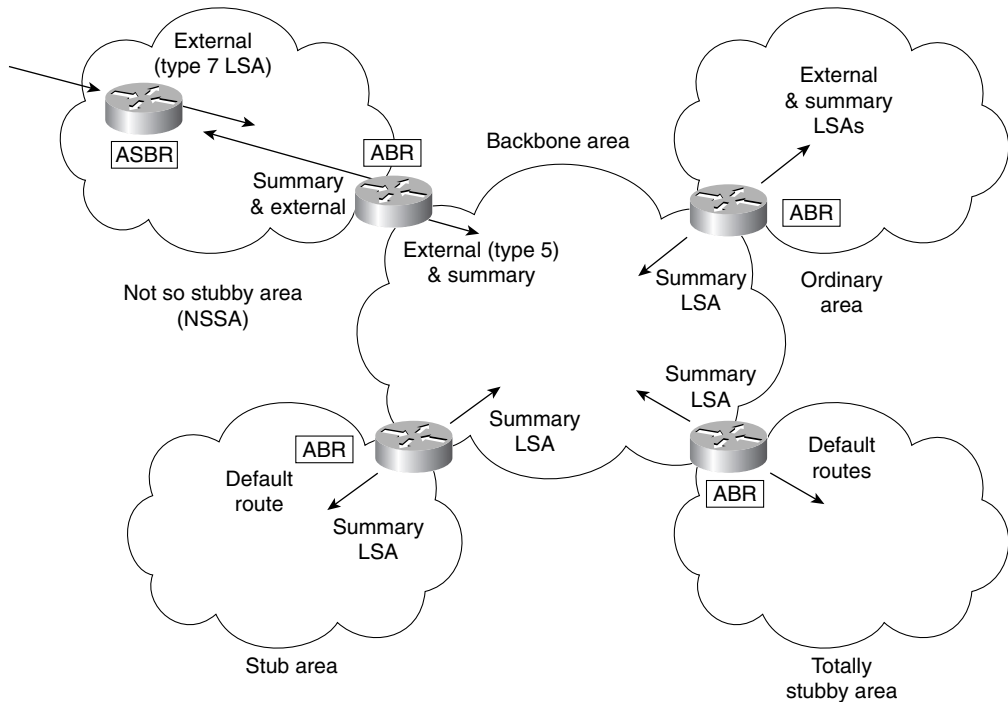
Figure 8-1 Router Definitions for OSPF

Figure 8-2 shows the connectivity and functionality of the different areas. The routers will send out routing updates and other network information through LSAs. The function or type of router will determine the LSAs that are sent.

Figure 8-2 The Different Types of OSPF Areas and LSA Propagation

Link-State Advertisements

Five commonly used types of link-state advertisements (LSAs) exist. Cisco uses six LSAs, which are briefly described here:

- **The router link LSA** — This LSA is generated for each area to which the router belongs. This LSA gives the link states to all other routers within an area. This LSA is flooded into an area. This is identified as a Type 1 LSA.
- **The network link LSA** — This LSA is sent out by the designated router and lists all the routers on the segment for which it is the designated router and has a neighbor relationship. The LSA is flooded to the whole area. This is identified as a Type 2 LSA.
- **The network summary link LSA** — This LSA is sent between areas and summarizes the IP networks from one area to another. It is generated by an ABR. This is identified as a Type 3 LSA.
- **The AS external ASBR summary link LSA** — This LSA is sent to a router that connects to the outside world (ASBR). It is sent from the ABR to the ASBR. The LSA contains the metric cost from the ABR to the ASBR. This is identified as a Type 4 LSA.
- **The external link LSA** — This LSA is originated by AS boundary routers and is flooded throughout the AS. Each external advertisement describes a route to a destination in another autonomous system. Default routes for the AS can also be described by AS external advertisements. This is identified as a Type 5 LSA.
- **The NSSA external LSA** — Identified as Type 7, these LSAs are created by the ASBR residing in a not so stubby area (NSSA). This LSA is similar to an autonomous system external LSA, except that this LSA is contained within the NSSA area and is not propagated into other areas, but it is converted into a Type 5 LSA by the ABR.

In the section “The ABRs and ASBR Propagation of LSAs,” Figure 8-3 shows the relationships between the different LSAs. This section discusses the router and network LSAs. The LSAs concerned with communication outside an area are considered later.

The Different Types of Areas

It is possible to create an OSPF network with only one area. This area is known as the backbone area or Area 0. In addition to the backbone area, which connects the other areas, OSPF networks use several other types of areas. The following are the different types of areas:

- **An ordinary or standard area** — This type of area connects to the backbone. The area is seen as an entity unto itself. Every router knows about every network in the area, and each router has the same topological database. However, the routing tables are unique from the perspective of the router and its position within the area.

- **A stub area** — This is an area that will not accept external summary routes. The LSA that is blocked is Type 5. The consequence is that the only way that a router within the stub area can see outside the autonomous system is by the use of a default route. Every router within the area can see every network within the area and the networks (summarized or not) within other areas. It is typically used in a hub-and-spoke network design.
- **A totally stubby area** — This area does not accept summary LSAs from the other areas or the external summary LSAs from outside the autonomous system. The LSAs blocked are Types 3, 4, and 5. The only way out of the totally stubby area is via a default route. A default route is indicated as the network 0.0.0.0. This type of area is particularly useful for remote sites that have few networks and limited connectivity with the rest of the network. This is a proprietary solution offered only by Cisco. Cisco recommends this solution if you have a totally Cisco shop because it keeps the topological databases and routing tables as small as possible.
- **An NSSA** — This area is used primarily to connect to ISPs, or when redistribution is required. In most respects, it is the same as the stub area. External routes are not propagated into or out of the area. It does not allow Type 4 or Type 5 LSAs. This area was designed as a special stub area for applications such as an area with a few stub networks but with a connection to a router that runs only RIP, or an area with its own connection to an Internet resource needed only by a certain division.

An NSSA is an area that is seen as a stub area but can receive external routes, which it will not propagate into the backbone area and thus the rest of the OSPF domain. Another LSA, Type 7, is created specifically for the NSSA. This LSA can be originated and communicated throughout the area, but it will not be propagated into other areas, including Area 0. If the information is to be propagated throughout the AS, it is translated into an LSA Type 5 at the NSSA ABR.

It is not always possible to design the network and determine where redistribution is to occur. RFC 1587, “The OSPF NSSA Option,” deals with this subject.

- **The backbone area** — This area is often referred to as Area 0, and it connects all the other areas. It can propagate all the LSAs except for LSA Type 7, which is translated into LSA Type 5 by the ABR.

Some restrictions govern creating a stub area or a totally stubby area. Because no external routes are allowed in these areas, the following restrictions are in place:

- No external routes are allowed.
- No virtual links are allowed.
- No redistribution is allowed.
- No ASBR routers are allowed.
- The area is not the backbone area.
- All the routers are configured to be stub routers.

The Operation of OSPF Across Multiple Areas

As you have learned so far in this chapter, there are many pieces to the puzzle of OSPF across multiple areas. Having identified the various pieces, you need to fit them together. Then you will see how the routing protocol operates across the various areas to maintain a coherent and accurate understanding of the autonomous system.

The ABRs and ASBR Propagation of LSAs

When a router is configured as an ABR, it generates summary LSAs and floods them into the backbone area. Routes generated within an area are Type 1 or Type 2, and these are injected as Type 3 summaries into the backbone. These summaries are then injected by the other ABRs into their own areas, unless they are configured as totally stubby areas. Any Type 3 or Type 4 LSA received from the backbone are forwarded into the area by the ABR.

The backbone also forwards external routes both ways unless the ABR is a stub router, in which case they are blocked.

If a summary is received from within the area, it cannot be forwarded. Summaries received from the backbone cannot be further summarized.

The flow and propagation of LSAs within and between areas is illustrated in Figure 8-3.

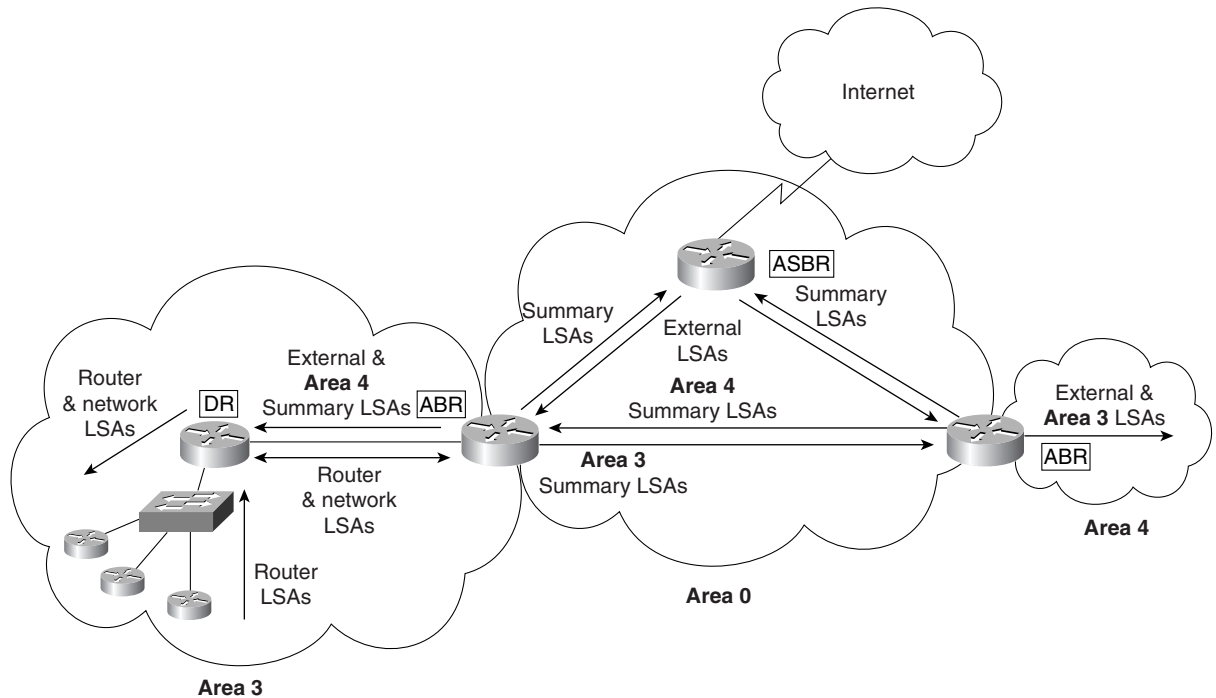
Certain conditions need to be met before any LSAs can be flooded out of all interfaces. The conditions that each interface must meet before an LSA can be transmitted out of that interface are given in the following list:

- The LSA was not received through the interface.
- The interface is in a state of exchange or full adjacency.
- The interface is not connected to a stub area (no LSA Type 5 will be flooded).
- The interface is not connected to a totally stubby area (no Type 3, 4, or 5 will be propagated).

OSPF Path Selection Between Areas

The OSPF routing table that exists on a router depends on the following factors:

- The position that the router has in the area and the status of the network
- The type of area in which the router is located
- Whether there are multiple areas in the domain
- Whether there are communications outside the autonomous system

Figure 8-3 *The Propagation of LSAs*

Remember the sequence of events: The router receives LSAs. It builds the topological database. Then it runs the Dijkstra algorithm, from which the shortest path is chosen and entered into the routing table. The routing table is therefore the conclusion of the decision-making process. It holds information on how that decision was made by including the metric for each link. This enables you to view the operation of the network.

Different LSAs are weighted differently in the decision-making process. It is preferable to take an internal route (within the area) to a remote network rather than to traverse multiple areas just to arrive at the same place. Not only does multiple-area traveling create unnecessary traffic, but it also can create a loop within the network.

The routing table reflects the network topology information and indicates where the remote network sits in relation to the local router.

The router will process the LSAs in this order:

1. The internal LSA (Type 1 and 2).
2. The LSAs of the AS (Type 3 and 4). If there is a route to the chosen network within the area (Type 1 or 2), this path will be kept.
3. The external LSAs (Type 5).

Calculating the Cost of a Path to Another Area

There are paths to networks in other areas, and then there are paths to networks in another autonomous system. The costs of these paths are calculated slightly differently.

The path to another area is calculated as the smallest cost to the ABR, added to the smallest cost to the backbone. Thus, if there were two paths from the ABR into the backbone, the shortest (lowest-cost) path would be added to the cost of the path to the ABR.

External routes are routes passed between a router within the OSPF domain and a router in another autonomous system or routing domain. The routes discovered by OSPF in this way can have the cost of the path calculated in one of two ways:

- **E1**—The cost of the path to the ASBR is added to the external cost to reach the next-hop router outside the AS.
- **E2**—The external cost of the path from the ASBR is all that is considered in the calculation. This is the default configuration. This is used when there is only one router advertising the route and no path selection is required. If both an E1 and an E2 path are offered to the remote network, the E1 path will be used.

At the side of the routing table is a column indicating the source of the routing information. Typically, this is the routing protocol. In the instance of OSPF, however, it includes the LSA type that provided the path.

Table 8-2 shows the codes used in the routing table.

Now that you understand the components and operation of multiple area OSPF, you should focus on some of the design implications of creating multiple areas, as described in the next section.

Table 8-2 *OSPF Routing Table Codes and Associated LSAs*

LSA Type	Routing Table Entry	Description
1 Router Link	O	This is generated by the router, listing all the links to which it is connected, their status, and their cost. It is propagated within the area.
2 Network Link	O	This is generated by the designated router on a multiaccess LAN to the area.
3 or 4 Summary Link (between areas)	O IA	LSA Type 3 includes the networks or subnets within an area that might have been summarized and that are sent into the backbone and between ABRs. LSA Type 4 is information sent to the ASBR from the ABR. These routes are not sent into totally stubby areas.
5 Summary Link/External Link (between autonomous systems)	O E1 or O E2	The routes in this LSA are external to the autonomous system. They can be configured to have one of two values. E1 will include the internal cost to the ASBR added to the external cost reported by the ASBR. E2 does not compute the internal cost—it just reports the external cost to the remote destination.

Design Considerations in Multiple Area OSPF

The major design consideration in OSPF is how to divide the areas. This is of interest because it impacts the addressing scheme for IP within the network.

An OSPF network works best with a hierarchical design, in which the movement of data from one area to another comprises only a subset of the traffic within the area itself.

It is important to remember that with all the interarea traffic disseminated by the backbone, any reduction of overhead through a solid hierarchical design and summarization is beneficial. The entire network benefits when fewer summary LSAs need to be forwarded into the backbone area. When network overhead is minimized, the network grows more easily.

With this in mind, summarization is the natural consequence. As shown in Chapter 2, “IP Addressing,” summarization is not something that can be imposed on a network. It must be part of the initial network design. The addressing scheme must be devised to support the use of summarization.

In designing any network, you need to consider the resources available and to make sure that none of these resources are overwhelmed, either initially or in the future. In the creation of areas, OSPF has tried to provide the means by which the network can grow without exceeding the available resources. However, this does not remove your responsibility as the network administrator to design a network that can run efficiently within the limits of the resources available. Cisco has laid down guidelines to help in the design of stable, responsive, and flexible OSPF networks.

It is also important in any design to allow for transitions or breaks in the network. OSPF has provided a cunning device called the *virtual link* that allows areas disconnected from the backbone area to appear directly connected to the backbone as required.

Finally, in any network design, you must consider the traditionally tricky topology of the WAN, in particular the nonbroadcast multiaccess (NBMA) connections that fall into neither one network topology nor another.

The following sections consider all of these subjects as they pertain to multiarea OSPF design.

Capacity Planning in OSPF

Although it is possible to have more than three areas (per router) in OSPF, the Cisco Technical Assistance Center (TAC) recommends that a greater number of areas be created only after careful consideration. The results of having more areas will vary depending on the router (memory and CPU), as well as network topology and how many LSAs are generated. The recommendation is not to exceed 50 routers in an OSPF area, but again, this is a guideline and not a strict rule. Remember that OSPF is very CPU-intensive in its maintenance of the databases and in the flooding of LSAs, as well as when it calculates the routing table, a process based on LSAs.

Therefore, it is not strictly the number of routers or areas that is important, but the number of routes and the stability of the network. You must consider these issues because the number of LSAs in your network is proportional to the amount of router resources required.

With this understanding, the general rules stated by Cisco for OSPF design are that the following numbers should not be exceeded:

- Routers per area: 50
- Neighbors per router: 60
- Areas per router: 3
- A router may not be a DR or BDR for more than 1 LAN

These are not hard and fast rules. The number of routers within an area depends on many factors; for example, a stub area with a 2500 router running over Ethernet is very different from area 0,

running 7500 routers over ATM. Some of the factors that influence the number of routers per area include the following:

- What type of area is it: stub, totally stub, or backbone? This determines the number of LSAs and how often and how much CPU and memory each SPF computation requires.
- What level of computing power do you have in the routers within the area? The smaller routers are not designed to manage large databases and to run the SPF algorithm continually.
- What kind of media do you have? The higher the bandwidth on the link, the less congestion within the router as it queues the packets for transmission.
- How stable is the network? How often LSAs will be propagated because of topology changes determines the need for bandwidth, CPU, and memory resources.
- If the area is running over NBMA, is the cloud fully meshed? To overcome the resources required to maintain a fully meshed network, Cisco suggests that a well-designed partial mesh over low-bandwidth links reduces the number of links and thus the amount of traffic and resources required.
- If the area has external connections, is there a large number of external LSAs? If the external connections are serviced with a default link, far less memory and CPU are required than if 500 external Internet links are propagated into the network.
- Do you have a hierarchical design with summarization? The greater the summarization, the smaller and fewer the LSA packets that need to be propagated.

Cisco states that, normally, a routing table with less than 500 KB could be accommodated with 2 to 4 MB RAM; large networks with greater than 500 KB might need 8 to 16 MB, or 32 to 64 MB if routes are injected from the Internet.

NOTE Further information is available on the Cisco web site at <http://www.cisco.com/warp/public/104/3.html#17.0> in the OSPF Design Guide.

The following sections describe how to determine the appropriate number of neighbors to which a router should be connected, or the number of areas to which an ABR should be connected. In designing a network, elements in the network that use resources, CPU, memory, and bandwidth must be evaluated and provided for, where appropriate. Luckily, Cisco has performed extensive tests to provide clear guidelines for the design and implementation of an OSPF network.

Number of Neighbors per Router

Increasing the number of neighbors increases the resources on the router that are allocated to managing those links. More importantly if there is a designated router (DR), the router that performs the DR function might become overloaded if there are many routers on the link. It might be advisable to select the DR through manual configuration to be the router with the most available

CPU and memory on the segment and to ensure that the router is not selected to be the DR for more than one link.

Number of Areas per ABR

For every area to which an ABR is connected, it will have a full topology table for that area. This could result in overloading the router before it has attempted to compute the best path. How many areas a router can support obviously depends on the caliber of the router and the size of the area. A good hierarchical design—where the maintenance of the areas is spread over a few routers—not only shares the resources, but also builds in a level of redundancy.

Summarization

One of the strengths of OSPF is the ability to scale the network. You can scale the network not only through the creation of multiple areas that limit the computation and propagation of routing updates, but also through the use of summarization. In Chapter 2, summarization was dealt with in great depth. This section builds on that knowledge and applies it to the design and implementation of multiarea OSPF.

In OSPF, two types of summarization exist:

- **Interarea summarization** —This is performed at the ABR and creates Type 3 and 4 LSAs.
- **External summarization** —This is performed at the ASBR and creates Type 5 LSAs.

Both have the same fundamental requirement of contiguous addressing.

OSPF is stringent in its demand for a solid hierarchical design, so much so that it has devised some commands to deal with situations that break its rules of structure.

The concept of the virtual link is explained in this section, while the commands with which to implement it are given in Chapter 9 in the section, “The **area virtual-link** Command.”

The Virtual Link

The main dictate in OSPF is that the multiple areas must all connect directly to the backbone area. The connection to the backbone area is through an ABR, which is resident in both areas and holds a full topological database for each area.

OSPF has provided a solution for the unhappy occasion when this rule cannot be followed. The solution is called a virtual link. If the new area cannot connect directly to the backbone area, a router is configured to connect to an area that does have direct connectivity.

The configuration commands create a tunnel to the ABR in the intermediary area. From the viewpoint of OSPF, the ABR has a direct connection.

The reasons such a situation might occur are as follows:

- There is no physical connection to Area 0. This might be because the organization has recently merged with another or because of a network failure.
- There are two Area 0s because of a network merger. These Area 0s are connected by another area (for example, Area 5).
- The area is critical to the company, and an extra link has been configured for redundancy.

Although the virtual link feature is extremely powerful, virtual links are not recommended as part of the design strategy for your network. Instead, they are a temporary solution to a connectivity problem. You must ensure that you observe the following when creating a virtual link:

- Both routers must share a common area.
- The areas involved cannot be stub areas.
- One of the routers must be connected to Area 0.

Figure 8-4 illustrates the use of a virtual link to provide a router in Area 10 connectivity to the backbone in Area 0.

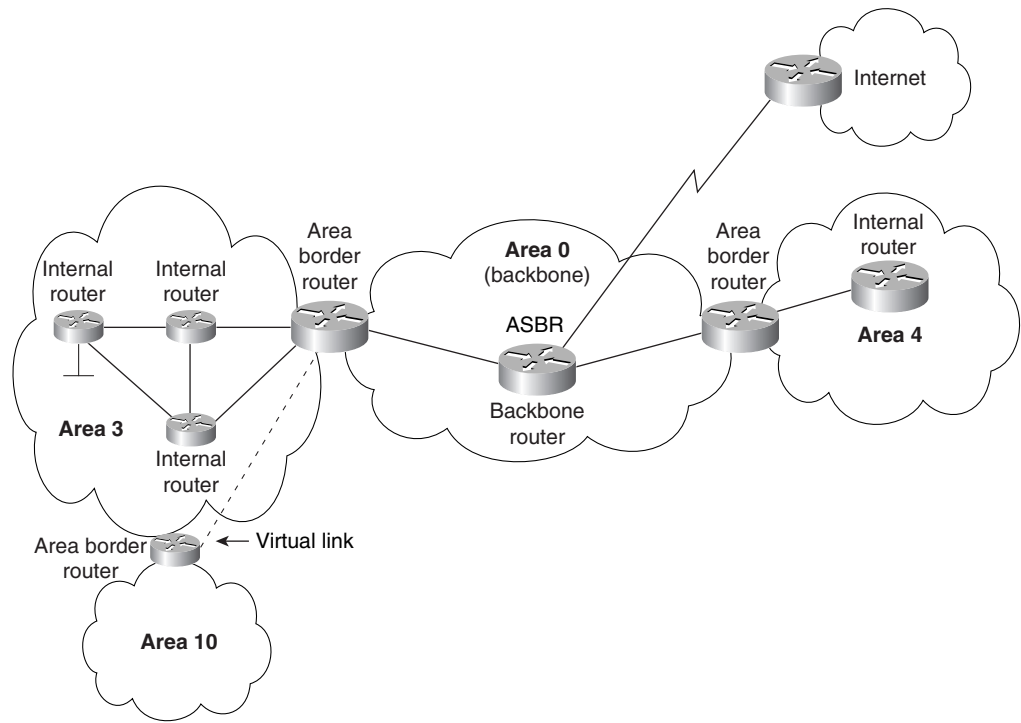
Multiple Area OSPF Over an NBMA Network

Another design consideration is the design of the NBMA network as part of the OSPF domain. There are two main ways to approach the inclusion of an NBMA network:

- The NBMA network can be defined as Area 0. The reasoning is that if the NBMA is used to connect all remote sites, all traffic will have to traverse this network. If the remote sites are made satellite areas, all traffic will have to traverse the NBMA, so it makes sense to make it the backbone area. This works well in a full-mesh environment, although it results in a large number of LSAs being flooded into the WAN and puts extra demands on the routers connecting to the NBMA network.
- In a hub-and-spoke NBMA network, it makes sense to assign the hub network as Area 0 with the other remote sites and the NBMA network as other areas. This is a good design if the satellite areas are stub areas because it means that the routing information—and, thus, network overhead—is kept to a minimum over the NBMA cloud. Depending on the design, the rest of the network might constitute one other area or multiple areas. This will depend on the size and growth expectations of the OSPF domain.

The configuration of a basic OSPF over an NBMA network is provided in Chapter 7.

Figure 8-4 Virtual Links in a Multiple Area OSPF Network



Foundation Summary

The “Foundation Summary” section of each chapter lists the most important facts from the chapter. Although this section does not list every fact from the chapter that will be on your exam, a well-prepared candidate should, at a minimum, know all the details in each “Foundation Summary” before going to take the exam.

The following list identifies the different OSPF routers:

- **Internal router** — Within an area, the functionality of the router is straightforward. It is responsible for maintaining a current and accurate database of every subnet within the area. It is also responsible for forwarding data to other networks by the shortest path. Flooding of routing updates is confined to the area.
- **Backbone router** — The design rules for OSPF require that all the areas be connected through a single area known as the *backbone area* or *Area 0*. A router within this area is referred to as a *backbone router*. It can also be an internal router, an ASBR or an ABR.
- **ABR**—This router is responsible for connecting two or more areas. It holds a full topological database for each area to which it is connected and sends LSA updates between the areas. These LSA updates are summary updates of the subnets within an area. It is at the area border that summarization should be configured for OSPF, because this is where the LSAs make use of the reduced routing updates to minimize the routing overhead on both the network and the routers.
- **ASBR**—To connect to the outside world, or to any other routing protocol, you need to leave the OSPF domain. OSPF is an *interior routing protocol* or *Interior Gateway Protocol (IGP)*; *gateway* is an older term for a router. The router configured for this duty is the ASBR. Although you can place this router anywhere in the OSPF hierarchical design, it should reside in the backbone area. Because any traffic leaving the OSPF domain is also likely to leave the router’s area, it makes sense to place the ASBR in a central location that all traffic leaving its area must traverse.

The five LSAs are as follows:

- **The router link** — This LSA states all the links to the router sending out the LSA. The list is of all the neighbors attached to the router. The LSA is flooded to the area.
- **The network link** — This LSA is sent out by the DR and lists all the routers on the segment for which it is the DR and has a neighbor relationship. The LSA is flooded to the whole area.
- **The network summary link** — This LSA is sent between areas and summarizes the IP networks from one area to another. It is generated by an ABR.

- **The AS external (ASBR) summary link** —This LSA is sent to a router that connects to the outside world (ASBR). It is sent from the ABR to the ASBR. The LSA contains the metric cost from the ABR to the ASBR.
- **The external link** —This LSA is originated by AS boundary routers and flooded throughout the AS. Each external advertisement describes a route to a destination in another autonomous system. Default routes for the AS can also be described by AS external advertisements.

Table 8-3 shows the codes used in the routing table.

Table 8-3 *OSPF Routing Table Codes and Associated LSAs*

LSA Type	Routing Table Entry	Description
1 Router Link	O	This is generated by the router, listing all the links to which it is connected, their status, and their cost. It is propagated within the area.
2 Network Link	O	This is generated by the DR on a multiaccess LAN to the area.
3 or 4 Summary Link (between areas)	O IA	LSA Type 3 includes the networks or subnets within an area that might have been summarized and that are sent into the backbone and between ABRs. LSA Type 4 is information sent to the ASBR from the ABR. These routes are not sent into totally stubby areas.
5 Summary Link/External Link (between autonomous systems)	O E1 or O E2	The routes in this LSA are external to the autonomous system. They can be configured to have one of two values. E1 will include the internal cost to the ASBR added to the external cost reported by the ASBR. E2 does not compute the internal cost—it just reports the external cost to the remote destination.

Some restrictions govern creating a stub area or a totally stubby area. Because no external routes are allowed in these areas, the following restrictions are in place:

- No external routes are allowed.
- No virtual links are allowed.
- No redistribution is allowed.
- No ASBR routers are allowed.
- The area is not the backbone area.
- All the routers are configured to be stub routers.

In designing an OSPF network, it is important to consider the following:

- Summarization:
 - Interarea: Performed at the ABR, creating type 3 and 4 LSAs
 - External: Performed at the ASBR creating type 5 LSAs
- Capacity planning:
 - Router per area: 50
 - Neighbors per router: 60
 - Areas per router: 3
 - A router may not be a DR or BDR for more than one LAN
- Virtual links: As a temporary solution during transition or after a break in the network
- NBMA networks:
 - Creating the NBMA network as area 0 if it is a fully meshed network used to connect all other sites
 - In a hub and spoke network, defining the hub as area 0, with the spokes forming other areas

Q&A

As mentioned in the introduction, “All About the CCNP, CCDP, and CCIP Certifications,” you have two choices for review questions. The questions that follow next give you a bigger challenge than the exam itself by using an open-ended question format. By reviewing now with this more difficult question format, you can exercise your memory better and prove your conceptual and factual knowledge of this chapter. The answers to these questions are found in Appendix A.

For more practice with examlike question formats, including questions using a router simulator and multichoice questions, use the exam engine on the CD.

1. In a totally stubby area, which routes are not propagated into the area?
2. Can a virtual link contain a stub area?
3. An ABR must be resident in which area?
4. What LSAs will the ABR forward?
5. State two advantages in creating areas in OSPF.
6. What is an external route, and on which type of router will this route be introduced?
7. Why is the use of summarization important in the design of OSPF?
8. How many routers does Cisco suggest is the limit to have in a single area?
9. What are the restrictions to be considered in the creation of a stub area or a totally stubby area?
10. A virtual link in OSPF is used to solve what problem?
11. State one disadvantage for making an NBMA cloud Area 0.
12. State one advantage in making the centralized routers and network resources dwell in Area 0 while the Frame Relay cloud and the stub remote LANs reside in satellite stub areas.
13. How does creating a number of areas in OSPF reduce the number of SPF calculations?
14. How does a stub area differ from the backbone area?
15. How does a totally stubby area differ from a stub area?
16. State the different LSA types.
17. Where does the backbone router reside, and what is its function?
18. There are two types of summarization. What are they?
19. For how many LANS does Cisco suggest a router should serve as a DR or a BDR?
20. Which router type creates LSA Types 3 and 4?

Scenarios

The following scenarios and questions are designed to draw together the content of the chapter and to exercise your understanding of the concepts. There is not necessarily a right answer. The thought process and practice in manipulating the concepts is the goal of this section. The answers to the scenario questions are found at the end of this chapter.

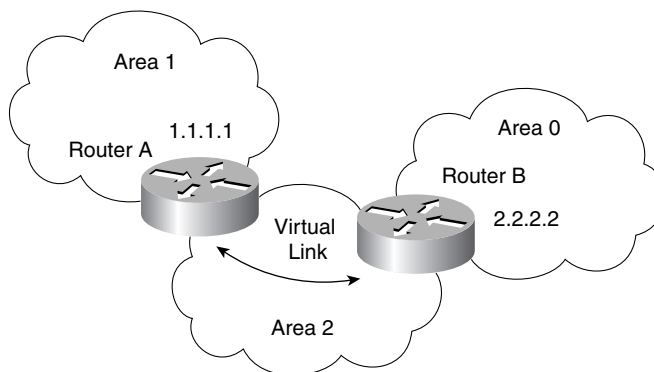
Scenario 8-1

A new network administrator recently joined the company and has found little documentation for the network. On drawing the topology of the network, the administrator has found a surprising configuration of a virtual link. Luckily, the administrator has come across such configurations in his previous job and understands their purpose.

After studying the figure provided, answer the following question.

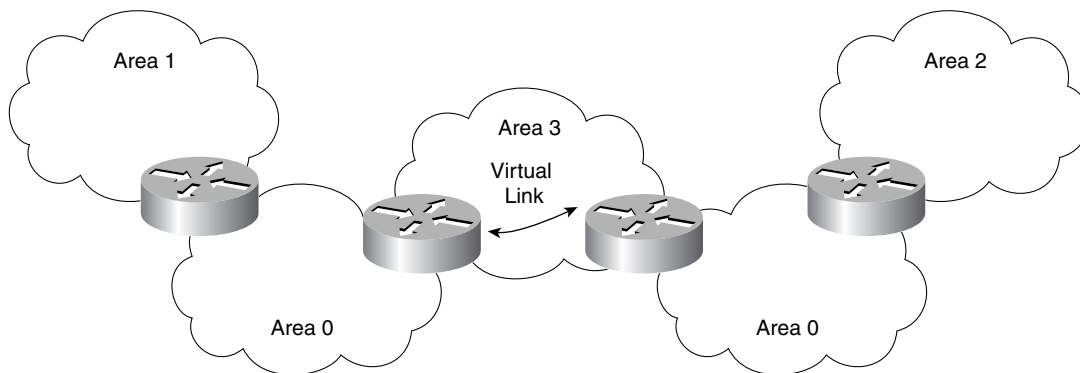
1. Explain the purpose of the virtual link in Figure 8-5.

Figure 8-5 *Network Diagram 1 for Scenario 8-1*



The following figure shows the network of another company for which the administrator worked previously. Examine the figure and answer the questions.

Figure 8-6 *Network Diagram 2 for Scenario 8-1*



2. Does the topology map in Figure 8-6 show a valid design?
3. Why would a company implement this design?

Scenario Answers

The answers provided in this section are not necessarily the only possible answers to the questions. The questions are designed to test your knowledge and to give practical exercise in certain key areas. This section is intended to test and exercise skills and concepts detailed in the body of this chapter.

If your answer is different, ask yourself whether it follows the tenets explained in the answers provided. Your answer is correct not if it matches the solution provided in the book, but rather if it has included the principles of design laid out in the chapter.

In this way, the testing provided in these scenarios is deeper: It examines not only your knowledge, but also your understanding and ability to apply that knowledge to problems.

If you do not get the correct answer, refer back to the text and review the subject tested. Be certain to also review your notes on the question to ensure that you understand the principles of the subject.

Scenario 8-1 Answers

1. *Explain the purpose of the virtual link in Figure 8-5.*

In this example, Area 1 does not have a direct physical connection into Area 0. A virtual link must be configured between Router A and Router B. Area 2 is to be used as a transit area, and Router B is the entry point into Area 0. This way, Router A and Area 1 will have a logical connection to the backbone.

2. *Does the topology map in Figure 8-6 show a valid design?*

Yes, the topology map in Figure 8-6 shows a valid design.

3. *Why would a company implement this design?*

OSPF allows for linking discontinuous parts of the backbone using a virtual link. In some cases, different Area 0s need to be linked together. This can occur, for example, if a company is trying to merge two separate OSPF networks into one network with a common Area 0. In other instances, virtual links are added for redundancy in case some router failure causes the backbone to be split in two. Whatever the reason might be, a virtual link can be configured between separate ABRs that touch Area 0 from each side and that have a common area between them.



This chapter covers the following topics, which you need to understand to pass the CCNP/CCDP/CCIP BSCI exam:

- Required configuration commands for a multiarea OSPF network
- Optional configuration commands for a multiarea OSPF network
- A working configuration of OSPF on a multiarea network
- Verifying the configuration of OSPF in a multiarea network
- Troubleshooting a multiarea OSPF network

Configuring OSPF Across Multiple Areas

The topics in this chapter describe the configuration of the routing protocol OSPF across multiple areas. This chapter assumes your knowledge of Chapter 6, “Using OSPF in a Single Area,” and Chapter 7, “Configuring OSPF in a Single Area,” which dealt with OSPF configuration in a single area, and Chapter 8, “Using OSPF Across Multiple Areas,” which introduced you to using OSPF in multiple areas. This chapter builds on your understanding and explains the configuration of OSPF over a large multiarea network. You will learn how to implement and manage an OSPF network in a real environment, one of multiple areas.

Some of the commands in this chapter are the same as those needed in the configuration of OSPF in a single area, in particular the required commands and the verification and troubleshooting commands. This chapter briefly reconsiders the OSPF commands required for configuration in order to put the multiarea commands in context. This chapter concentrates on the configuration of a multiarea environment.

“Do I Know This Already?” Quiz

The purpose of the “Do I Know This Already?” quiz is to help you decide what parts of this chapter to use. If you already intend to read the entire chapter, you do not necessarily need to answer these questions now.

The 13-question quiz, derived from the major sections in the “Foundation Topics” portion of the chapter, helps you to determine how to spend your limited study time.

Table 9-1 outlines the major topics discussed in this chapter and the “Do I Know This Already?” quiz questions that correspond to those topics.

Table 9-1 “Do I Know This Already?” Foundation Topics Section-to-Question Mapping

Foundation Topics Section	Questions Covered in This Section
Required Configuration Commands for a Multiarea OSPF Network	1–3
Optional Configuration Commands for a Multiarea OSPF Network	4–6
Verifying the Configuration of OSPF in a Multiarea Network	7–10
Troubleshooting a Multiarea OSPF Network	11–13

CAUTION The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark this question wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

- Which of the following is the correct command for configuring OSPF?
 - Router(config)# **router ospf** *process-number*
 - Router(config)#**router ospf** *router-ID*
 - Router(config)#**router ospf** *autonomous-system-number*
 - Router(config-if)#**router ospf** *process-id*
- Which of the following commands is the correct command for placing the subnet 172.16.20.128 in Area 1, and all other subnets within the classful address in Area 0?
 - network 172.16.20.128 0.0.0.0 area 1**
network 0.0.0.0 255.255.255.255 area 0
 - network 172.16.20.128 0.0.0.255 area 1**
network 172.16.0.0 0.0.255.255 area 0
 - network 172.16.20.128 0.0.0.0 area 1**
network 172.16.20.0 255.255.255.255 area 0
 - network 172.16.20.128 0.0.0.7 area 1**
network 172.16.20.0 255.255.255.255 area 0

3. Which of the following statements are true about placing subnets into different areas?
 - a. The **network** command used to assign interfaces and subnets to different areas can be configured on any OSPF router.
 - b. The **network** command is used at the interface level to assign interfaces and their subnets into different areas.
 - c. The **network** command uses the subnet mask to identify the portion of the address to consider in assigning areas.
 - d. The **network** command uses a wildcard mask to identify the portion of the address to consider in assigning areas.
4. Which of the following statements are true of the command **summary-address 172.16.20.0 255.255.255.0** ?
 - a. This command would be configured on an ASBR.
 - b. This command is used to summarize routes between areas.
 - c. This command is used to summarize addresses advertised to the outside world.
 - d. This command is configured on ABR routers.
5. Which command would be used on an ABR to configure a cost of a default route propagated into a stub or NSSA area?
 - a. **default-cost 30**
 - b. **area 1 default-cost 30**
 - c. **ip ospf default cost 30**
 - d. **ip ospf area 1 default-cost 30**
6. What is the purpose of the command **area 3 stub no-summary** ?
 - a. To create Area 3 as a totally stubby area
 - b. To create Area 3 as a stub area
 - c. To create Area 3 as an NSSA area
 - d. To restrict summaries passing between areas
7. Which of the following commands shows the correct syntax for defining a virtual link?
 - a. **area 1 virtual-link 10.10.10.10**
 - b. **ip ospf area 1 virtual-link 10.10.10.10**

- c. **ip ospf virtual-link area 1**
 - d. **area 1 virtual-link 10.10.10.0 255.255.255.0 172.16.20.131 255.255.255.0**

- 8. Which command is used to show the OSPF ABRs and ASBRs for which the internal router has entries in its routing table?
 - a. **show ip ospf external routers**
 - b. **show ip ospf route**
 - c. **show ip ospf border database**
 - d. **show ip ospf border-routers**

- 9. Which of the following is the correct command to show virtual links?
 - a. **show virtual-links**
 - b. **show ip ospf virtual-links**
 - c. **show ip ospf links**
 - d. **show ip ospf neighbors**

- 10. Which of the following are shown in the **show ip ospf database** command?
 - a. Router ID number
 - b. Hello timer intervals
 - c. Advertising router ID
 - d. Number of interfaces detected for router

- 11. If a router configured for OSPF routing is not seeing an OSPF neighbor on an attached network, which of the following actions should be taken?
 - a. Ensure that the OSPF autonomous system number is configured correctly on both ends
 - b. Make sure that both routers are configured with the same IP mask, Interface Hello timer, OSPF Hello interval, and OSPF dead interval
 - c. Make sure that both neighbors are part of the same area and area type
 - d. Ensure the virtual links are configured in the same subnet

- 12. Which of the following events are shown the **debug ip ospf events** command?
 - a. Adjacencies
 - b. Designated router selection

- c. Virtual links
 - d. SPF calculations
13. Which of the following commands is useful for analyzing the messages traveling between the local and remote hosts?
- a. **show ip ospf traffic**
 - b. **debug ip ospf data**
 - c. **debug ip packet**
 - d. **show ip ospf retransmission-list**

The answers to the "Do I Know This Already?" quiz are found in Appendix A, "Answers to Chapter 'Do I Know This Already?' Quizzes and Q&A Sections." The suggested choices for your next step are as follows:

- **6 or less overall score** —Read the entire chapter. This includes the "Foundation Topics" and "Foundation Summary" sections, the "Q&A" section, and the "Scenarios" at the end of the chapter.
- **7–9 overall score** —Begin with the "Foundation Summary" section, and then go to the "Q&A" section and the "Scenarios" at the end of the chapter. If you have trouble with these exercises, read the appropriate sections in "Foundation Topics."
- **10 or more overall score** —If you want more review on these topics, skip to the "Foundation Summary" section, and then go to the "Q&A" section and the "Scenarios" at the end of the chapter. Otherwise, move to the next chapter.

Foundation Topics

Required Configuration Commands for a Multiarea OSPF Network

As you saw in Chapter 7, it is necessary to start the routing process on the router in order to define how it is to operate. As part of starting the process, you must tell the process how to run. Many OSPF commands are optional, as described in the section “Optional Configuration Commands for a Multiarea OSPF Network,” but the following need to be defined at the startup of the process:

- **Participating router interfaces** —The router might not want to have all its interfaces send or receive OSPF routing updates. A classic example is a dialup line to a remote office. If there is only one subnet at the remote office, it would be more efficient to use default and static route commands, because any updates would dial the line.
- **Identification of the area** —The router will define which area it is in on a per-interface basis.
- **A router ID** —This allows the router to be uniquely identified by a single address. The ID of the router advertising a link can be used to determine the next logical hop, for example, if that link is used in the path selection to a remote network.

The following commands, which you will learn more about in the next sections, are necessary to configure a multiarea OSPF network:

- The OSPF **router** command
- The OSPF **network** command

Enabling the OSPF Routing Protocol

When configuring the router for the first time, there is no IP routing protocol running on the Cisco router (unless the SETUP script is used). This is not true of other protocols, however (refer to Chapter 7).

To configure OSPF as the routing protocol, use the following command:

```
Router(config)#router ospf process-number
```

Recall that *process-number* is a number local to the router. It is possible to have more than one process running on a router, although this is an unusual and expensive configuration in terms of router resources. The process number does not have to be the same on every router in the area or the

autonomous system. In the interest of sanity, however, many administrators assign the same number to the routers.

NOTE A common error in configuration is to confuse the process ID with the router ID or the area ID. These are not related in any way. The process ID is simply a mechanism to allow more than one process to be configured on a router. The router ID is the mechanism by which a router is identified within the OSPF domain, and the area ID is a mechanism of grouping routers that share full knowledge of OSPF-derived routes within the OSPF area.

Enabling the **network** Command

The **network** command was explained in Chapter 7 in terms of identifying the interfaces that participated in the OSPF routing process. In this chapter, the **network** command is used to identify not only the interfaces that are sending and receiving OSPF updates, but also the area in which they reside. Defining the areas with the **network** command is used on an area border router (ABR).

The following is the syntax for the OSPF **network** command:

```
Router(config-router)#network network-number wildcard-mask area area-number
```

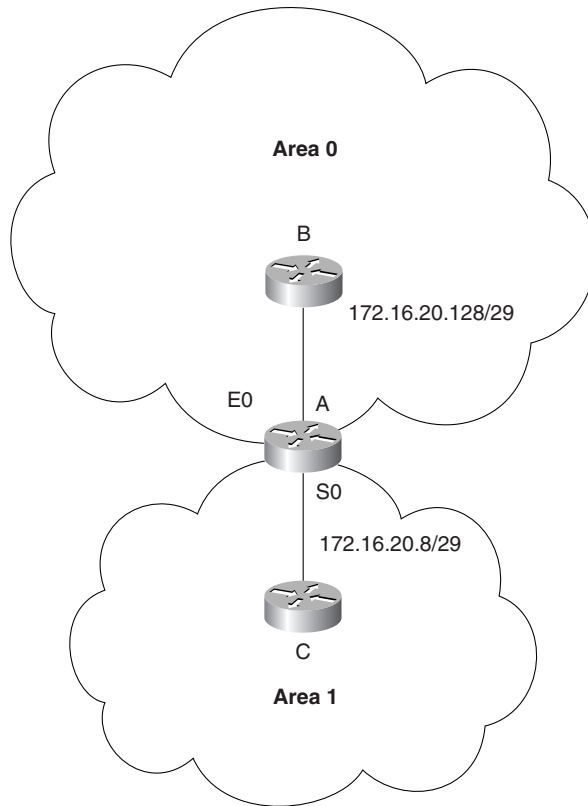
NOTE The area requested in the preceding syntax is the area in which the interface or interfaces configured with the network address reside.

You must take care in the use of the wildcard mask. In a single-area configuration, all the interfaces are in the same area. The **network** commands just identify the network numbers in use. Therefore, they can be configured to the classful network address provided by the IANA, as they are in IGRP and RIP. The only reason to be more specific would be to exclude some interfaces from the OSPF domain.

Example 9-1 shows two interfaces, each with a subnet of the same major network where the interfaces are placed into different areas. The network number has been subnetted into the last octet so that you can truly see the power and granularity of the wildcard mask at work. Figure 9-1 illustrates this configuration.

Example 9-1 The **network** Command for Router A

```
RouterA(config)#router ospf 100  
RouterA(config-router)#network 172.16.20.128 0.0.0.7 area 0  
RouterA(config-router)#network 172.16.20.8 0.0.0.7 area 1
```

Figure 9-1 The `network` Command

Optional Configuration Commands for a Multiarea OSPF Network

The word *optional* is used here to mean not absolutely necessary, implying that OSPF will run without the optional configuration commands. This does not mean that OSPF runs well or efficiently without them, however. A few of the OSPF commands, optional or not, are necessary in the configuration of an efficient multiarea OSPF network. The following list shows those optional OSPF commands that are important to the maintenance of an efficient network:

- The **area range** command configured on an ABR
- The **summary-address** command for use on an autonomous system border router (ASBR)
- The **area area-id stub** command to define a stub area
- The **area area-id stub no-summary** command to define a totally stubby area

- The **area default-cost** command for determining the cost of default routes that enter the area
- The **area virtual-link** commands used to create a virtual link

The area range Command

The **area range** command is configured on an ABR because it dictates the networks that will be advertised out of the area.

Use the **area** router configuration command with the **range** keyword to consolidate and summarize routes at an area boundary. This reduces the size of the databases, which is particularly useful in the backbone area because it is the transit area for all the other areas. Use the **no** form of this command to disable this function for the specified area:

```
Router(config-router)#area area-id range address mask
Router(config-router)#no area area-id range address mask
```

In the preceding syntax, *area-id* is the identifier (ID) of the area about which routes are to be summarized. It can be specified as either a decimal value or an IP address. Here, *address* is the IP address, and *mask* is the IP mask.

Example 9-2 shows the configuration required to summarize the following five individual subnets (which can address six hosts each) into one subnet. This summarized subnet can then be propagated across the OSPF network, saving both bandwidth and CPU:

- 172.16.20.128/29
- 172.16.20.144/29
- 172.16.20.160/29
- 172.16.20.176/29

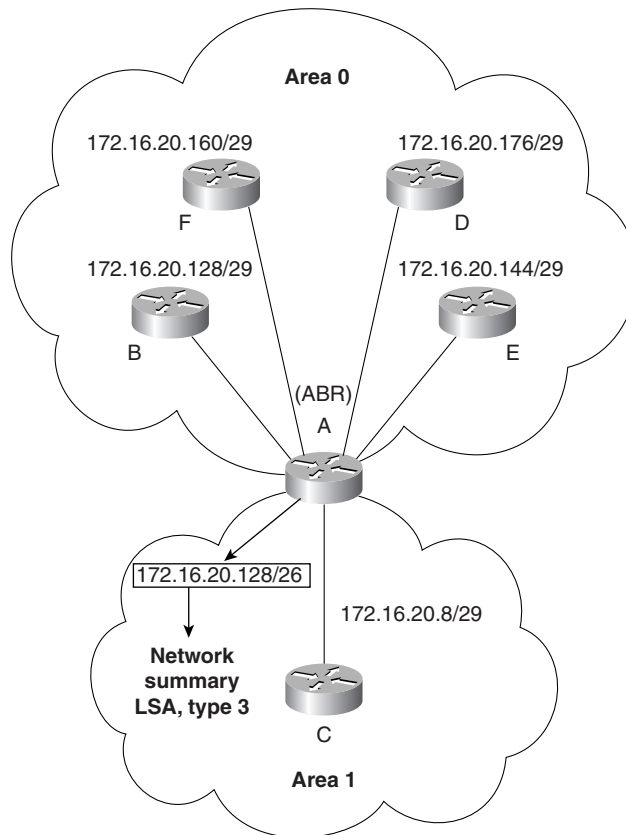
These subnets are summarized into one subnet:

- 172.16.20.128/26

This one subnet will then be propagated into Area 1 (see Figure 9-2).

Example 9-2 The OSPF area range Command for an ABR

```
RouterA(config)#router ospf 100
RouterA(config-router)#network 172.16.20.128 0.0.0.7 area 0
RouterA(config-router)#network 172.16.20.144 0.0.0.7 area 0
RouterA(config-router)#network 172.16.20.176 0.0.0.7 area 0
RouterA(config-router)#network 172.16.20.160 0.0.0.7 area 0
RouterA(config-router)#network 172.16.20.8 0.0.0.7 area 1
RouterA(config-router)#area 0 range 172.16.20.128 255.255.255.192
```

Figure 9-2 *The OSPF area range Command for an ABR*

NOTE The area ID requested is the area from which the subnets originated. It is not the destination area.

The summary-address Command

The **summary-address** command is used on the ASBR to summarize the networks to be advertised to the outside world. Routes received into OSPF via redistribution from another routing protocol are seen as external routes. The **summary-address** command allows the summarization of routes to be redistributed into OSPF from other routing protocols.

The syntax for the OSPF **summary-address** command for an ASBR is as follows:

```
Router(config-router)#summary-address address mask [not-advertise][tag tag]
```

In the preceding syntax, *address* is the summary address designated for a range of addresses, and *mask* is the IP subnet mask used for the summary route.

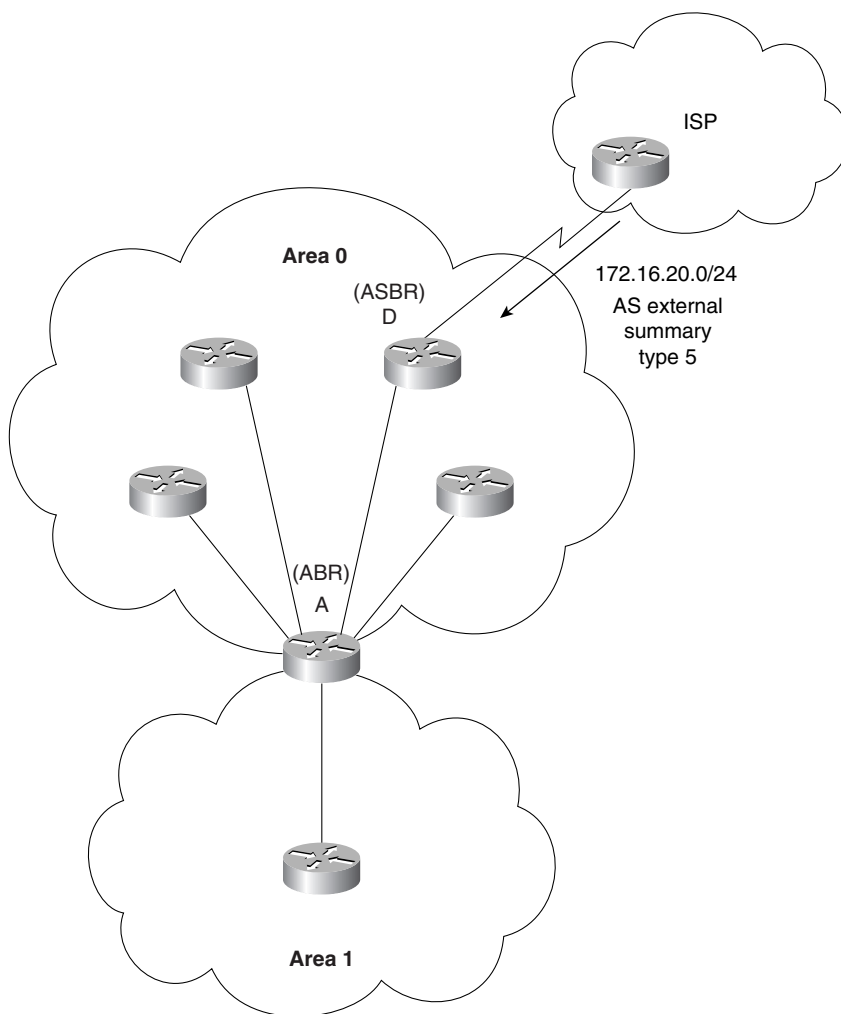
The design and implementation of the addressing scheme are crucial to the success of the OSPF network and cannot be stressed too strongly. Refer to Chapter 2, “IP Addressing,” for information about summarization and Chapter 3, “Designing IP Networks,” for details on IP addressing.

Example 9-3 shows the summarization of the network address 172.16.20.0, received from the ISP and propagated throughout the autonomous system. Figure 9-3 illustrates this example.

Example 9-3 *The OSPF summary-address Command for an ASBR*

```
RouterD(config)#router ospf 100
RouterD(config-router)#network 172.16.20.176 0.0.0.7 area 0
RouterD(config-router)#summary-address 172.16.20.0 255.255.255.0
```

Figure 9-3 *The OSPF summary-address Command for an ASBR*



The area *area-id* stub Command

After designing the addressing scheme for the network, it should be clear which areas, if any, are suitable candidates for configuration as stub, totally stubby, or not so stubby areas.

The syntax for the OSPF router command for a stub area is as follows:

```
Router(config-router)#area area-id stub
```

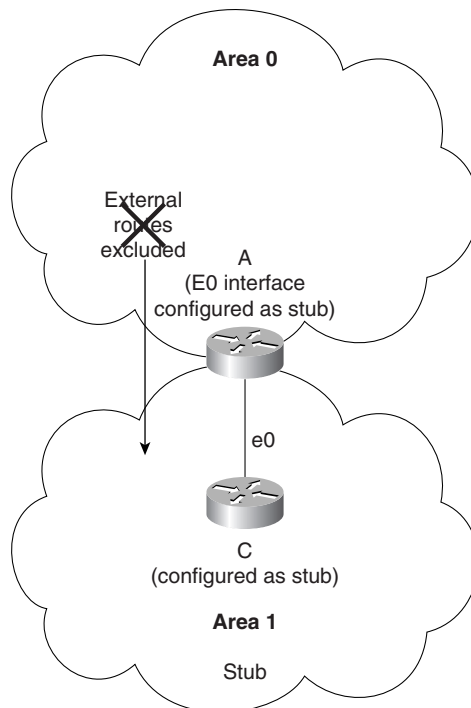
Example 9-4 shows the creation of a stub area. Note that both the ABR and the internal router share the stub area configuration. Figure 9-4 illustrates this example.

Example 9-4 The Configuration of a Stub Area

```
RouterC(config)#router ospf 100
RouterC(config-router)#network 0.0.0.0 255.255.255.255 area 1
RouterC(config-router)#area 1 stub

RouterA(config)#router ospf 100
RouterA(config-router)#network 172.16.20.128 0.0.0.7 area 0
RouterA(config-router)#network 172.16.20.8 0.0.0.7 area 1
RouterA(config-router)#area 0 range 172.16.20.128 255.255.255.192
RouterA(config-router)#area 1 stub
```

Figure 9-4 The Configuration of a Stub Area



NOTE All OSPF routers inside a stub area must be configured as stub routers. Whenever an area is configured as a stub, all interfaces that belong to that area will start exchanging Hello packets with a flag that indicates that the interface is a stub. The flag is a bit in the Hello packet called the E bit. The E bit is set to 0. All routers that have a common segment must agree on this flag. If they do not, they will not become neighbors, and routing will not take effect.

The `area area-id stub no-summary` Command

The syntax for the OSPF command for a totally stubby area is as follows:

```
Router(config-router)#area area-id stub no-summary
```

This addition of the **no-summary** parameter informs the ABR not to send summary updates from other areas into the area. This command needs to be configured only on the ABR because it is the only router with this responsibility. This command is configurable only on a Cisco router because it is a proprietary feature. All the other routers are configured as stub-area internal routers.

Example 9-5 shows the configuration of a totally stubby area. Figure 9-5 illustrates this example.

Example 9-5 *The Configuration of a Totally Stubby Area*

```
RouterE(config)#router ospf 100
RouterE(config-router)#network 172.16.20.144 0.0.0.192 area 0
RouterE(config-router)#network 172.16.20.16 0.0.0.7 area 2
RouterE(config-router)#area 2 stub no-summary

RouterG(config)#router ospf 100
RouterG(config-router)#network 0.0.0.0 255.255.255.255 area 2
RouterG(config-router)#area 2 stub
```

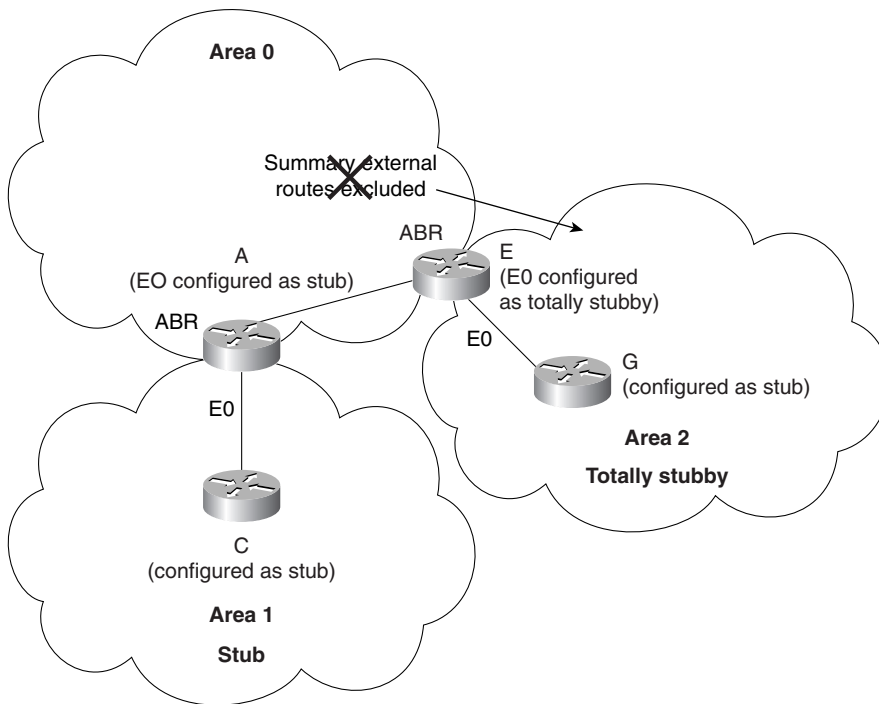
As a totally stubby area, no summary or external routes are propagated by the ABR into the area. To reach networks and hosts outside their area, a workstation must send to a default route, which the ABR advertises into the area.

The `area default-cost` Command

To define the cost to the default route, the following command is used. If the cost is not specified, the path will be calculated as the internal area cost plus 1:

```
Router(config-router)#area area-id default-cost cost
```

The ABR attached to the stub area automatically generates and advertises a default route with a destination of 0.0.0.0 into the stub area.

Figure 9-5 *The Configuration of a Totally Stubby Area***Example 9-6** *The OSPF Command for the Default Route Propagated into the Area*

```

RouterC(config-router)#router ospf 100
RouterC(config-router)#network 0.0.0.0 255.255.255.255 area 1
RouterC(config-router)#area 1 stub

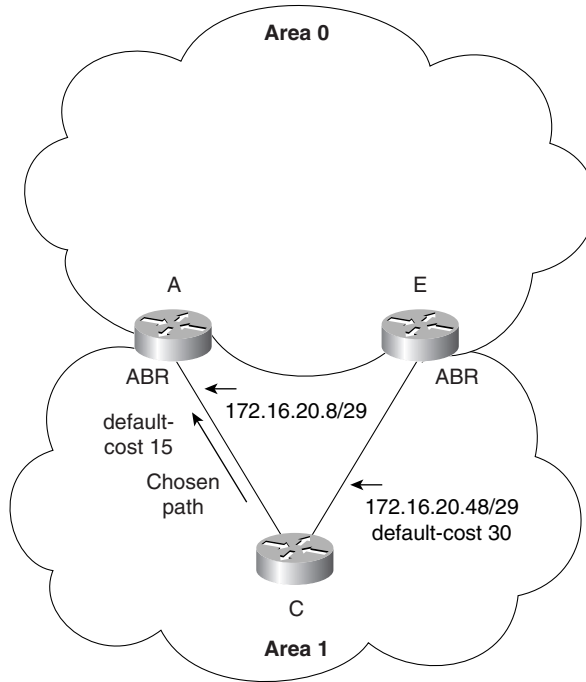
RouterA(config-router)#router ospf 100
RouterA(config-router)#network 172.16.20.128 0.0.0.7 area 0
RouterA(config-router)#network 172.16.20.8 0.0.0.7 area 1
RouterA(config-router)#area 0 range 172.16.20.128 255.255.255.192
RouterA(config-router)#area 1 stub
RouterA(config-router)#area 1 default-cost 15

RouterE(config-router)#router ospf 100
RouterE(config-router)#network 172.16.20.144 0.0.0.7 area 0
RouterE(config-router)#area 1 stub
RouterE(config-router)#area 1 default-cost 30
RouterE(config-router)#area 0 range 172.16.20.128 255.255.255.192

```

Example 9-6 shows how the default cost can be set in the configuration. Setting a cost on the default route is useful when the stub area has more than one ABR. The default cost allows the ABR used to exit the area. Figure 9-6 illustrates this example.

Figure 9-6 *The OSPF Command for the Default Route Propagated into the Area*



You need to configure the **area default-cost** command only on the ABR because it is the only router with this responsibility. However, Example 9-6 shows the configuration on both routers to illustrate the choice. The second ABR, Router E, will only be used if Router A fails. If there were no configuration on Router A, it would still be used by all internal routers as the ABR because the default cost is 1.

The area virtual-link Command

When it is not possible to connect an area to Area 0 directly, one solution is to create a tunnel called a *virtual link*. This is remarkably easy to configure. As with many things in OSPF, of course, this ease of configuration belies the complexity of the technology being used. Many things can go wrong. The most common problem is in the address of the other end of the virtual link. The

command **area virtual-link** is configured between ABRs that share a common area; at least one of the ABRs must be in Area 0. The command, issued at both ABRs, states the transit area and the router ID of the remote destination ABR. This creates essentially a tunnel through the transit area, which, although it might involve many routers to forward the traffic, appears to the remote ABRs as next hops.

NOTE The **area area-id virtual-link** command might be included in the BSCI exam and, for that reason, is worth mentioning. In practice, virtual links are a design nightmare and are best avoided. They are useful when mending a network on a temporary basis while awaiting a moment's peace to rectify the design of the network.

The syntax to configure a virtual link is as follows:

```
Router(config-router)#area area-id virtual-link router-id
```

Here, *area-id* is the ID assigned to the transit area for the virtual link. In addition, *router-id* is the router ID of the virtual link neighbor.

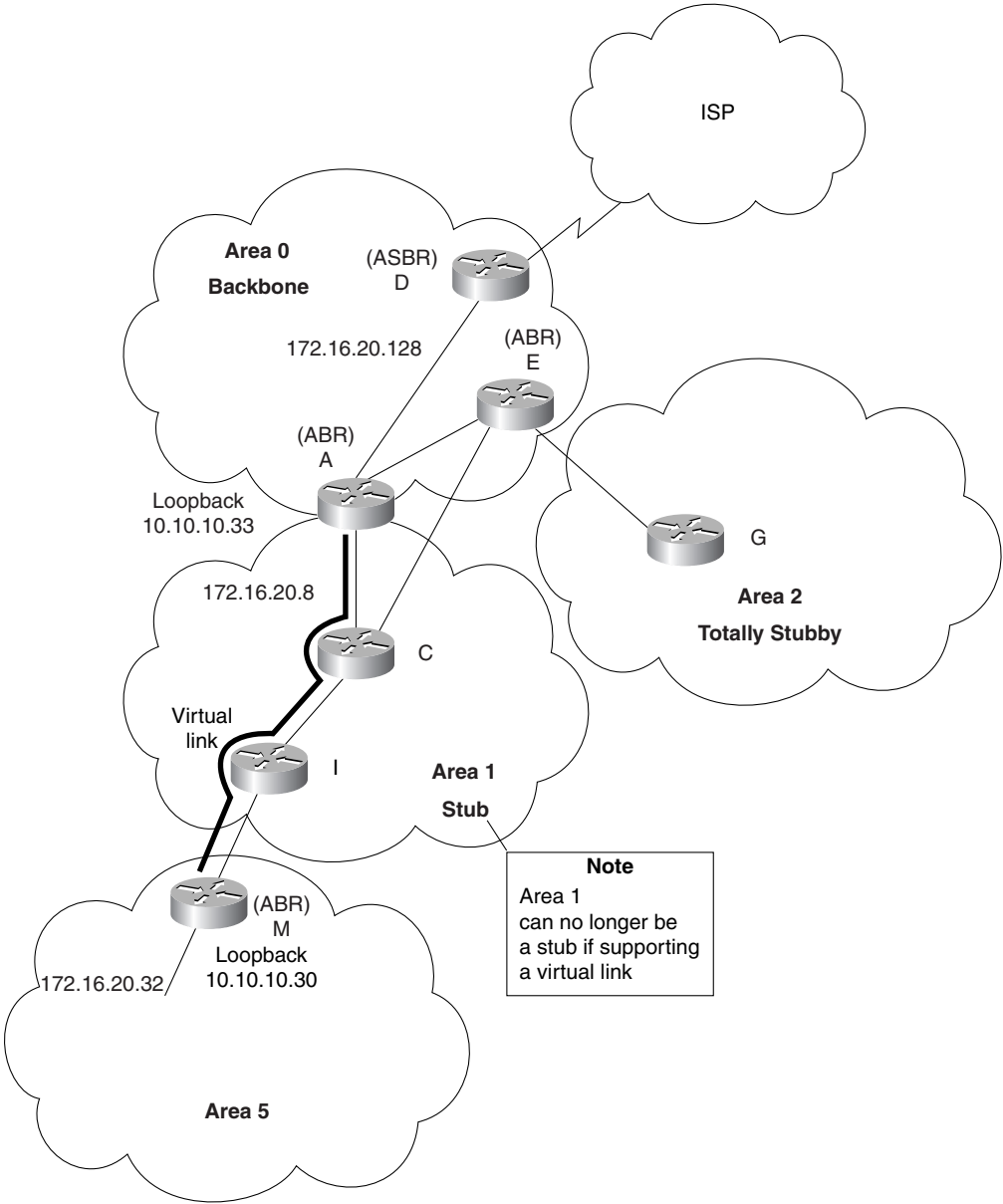
Example 9-7 shows the setting of the loopback interfaces that provide the router ID. It then shows the configuration of the virtual link through the network. Figure 9-7 illustrates this example.

Example 9-7 Configuring a Virtual Link

```
RouterA(config)#interface loopback 0
RouterA(config-if)#ip address 10.10.10.33 255.255.255.255
RouterA(config)#router ospf 100
RouterA(config-router)#network 172.16.20.128 0.0.0.7 area 0
RouterA(config-router)#network 10.10.10.33 0.0.0.0 area 0
RouterA(config-router)#area 0 range 172.16.20.128 255.255.255.192
RouterA(config-router)#area 1 default-cost 15
RouterA(config-router)#area 1 virtual-link 10.10.10.30

RouterM(config)#loopback interface 0
RouterM(config-if)#ip address 10.10.10.30 255.255.255.255
RouterM(config)#router ospf 100
RouterM(config-router)#network 172.16.20.32 0.0.0.7 area 5
RouterM(config-router)#network 10.10.10.30 0.0.0.0 area 0
RouterM(config-router)#area 1 virtual-link 10.10.10.33
```

Figure 9-7 *Configuring a Virtual Link*



A Working Configuration of OSPF on a Multiarea Network

Example 9-8 is a working configuration tested for verification. It includes many of the commands explained earlier in this chapter. Here you see an entire working configuration rather than the relevant segment for configuring a particular networking nuance. Use the configuration in conjunction with Figure 9-8.

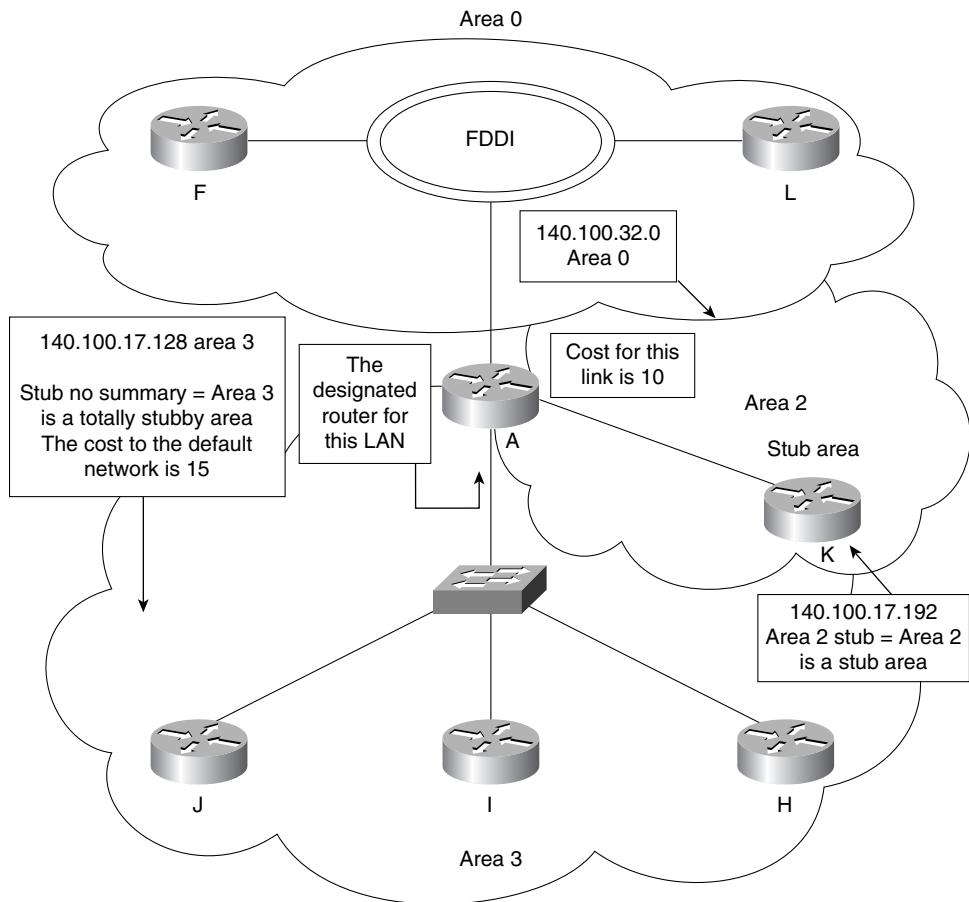
Example 9-8 *Configuring OSPF in a Multiarea Network on Router A*

```
RouterA(config)#router ospf 100
RouterA(config-router)#network 140.100.17.128 0.0.0.15 area 3
RouterA(config-router)#network 140.100.17.192 0.0.0.15 area 2
RouterA(config-router)#network 140.100.32.0 0.0.0.255 area 0
RouterA(config-router)#area 2 stub
RouterA(config-router)#area 3 stub no-summary
RouterA(config-router)#area 3 default-cost 15
!
RouterA(config-router)#interface FastEthernet0
RouterA(config-if)#ip address 140.100.17.129 255.255.255.240
RouterA(config-if)#no ip directed-broadcast
RouterA(config-if)#ip ospf priority 100
!
RouterA(config-if)#interface FastEthernet1
RouterA(config-if)#ip address 140.100.17.193 255.255.255.240
RouterA(config-if)#no ip directed-broadcast
RouterA(config-if)#ip ospf cost 10
!
RouterA(config-if)#interface Fddi0
RouterA(config-if)#ip address 140.100.32.10 255.255.255.240
RouterA(config-if)#no ip directed-broadcast
RouterA(config-if)#no keepalive

RouterA(config-if)#exit
```

CAUTION The **ip directed-broadcasts** command allows broadcasts to be sent across the network as a unicast and to be released as a broadcast on the interface for which it is destined. Although this is a very useful option, it has been exploited by hackers and saboteurs. It is therefore recommended that you disable this feature unless it is specifically required.

It is imperative to check any configuration on a network device because any errors could potentially bring down the entire network. To verify the configuration, a wealth of Cisco commands are available. They are covered in the following section.

Figure 9-8 *Example 9-8 Network*

Verifying the Configuration of OSPF in a Multiarea Network

The **show** commands shown here are in addition to the commands described in Chapter 7 in the section “Checking the Configuration of OSPF on a Single Router.” The single router commands are also extremely useful in a multiarea configuration. They are all invaluable in the configuration, troubleshooting, and maintenance of a live network. The following are additional commands that you can use in conjunction with single router commands when verifying OSPF operation on a multiarea network:

- The **show ip ospf border-routers** command
- The **show ip route** command

- The `show ip ospf database` command
- The `show ip ospf virtual-links` command

The `show ip ospf border-routers` Command

The following command shows the OSPF ABRs and ASBRs for which the internal router has entries in its routing table. This command is excellent for troubleshooting configuration errors and understanding how the network is communicating about its routes.

```
Router#show ip ospf border-routers
```

Example 9-9 shows the output of this command.

Example 9-9 The `show ip ospf border-routers` Output

Router# <code>show ip ospf border-routers</code>							
OSPF Process 100 internal Routing Table							
Destination	Next Hop	Cost	Type	Rte Type	Area	SPF No	
160.89.97.53	144.144.1.53	10	ABR	INTRA	0.0.0.3	3	
160.89.103.51	160.89.96.51	10	ABR	INTRA	0.0.0.3	3	
160.89.103.52	160.89.96.51	20	ASBR	INTER	0.0.0.3	3	
160.89.103.52	144.144.1.53	22	ASBR	INTER	0.0.0.3	3	

Table 9-2 explains the meaning of the important fields in the output of the `show ip ospf border-routers` command.

The `show ip ospf border-routers` command is useful to verify that the configuration has worked and that the OSPF network is functioning correctly. In a multiarea network, the `show ip ospf border-routers` command can immediately indicate why users cannot connect outside their area. This command is helpful in extracting information about the border routers from what could be a long routing table, within which this information is scattered.

Table 9-2 Explanation of the `show ip ospf border-routers` Command Output

Field	Explanation
OSPF Process 100 internal Routing Table	This is the OSPF routing process ID for the router.
Destination	This is the router ID of the destination router, whether an ABR or an ASBR.
Next Hop	If the ABR or ASBR is not directly connected, this is the address of the next logical hop in the chosen path to the ABR or ASBR.
Cost	This is the metric or cost of taking this path to the destination.

Table 9-2 *Explanation of the show ip ospf border-routers Command Output (Continued)*

Field	Explanation
Type	This states whether the destination router is an ABR, an ASBR, or both.
Rte Type	This is the type of route; it is either an intra-area or interarea route.
Area	This is the area ID of the area from which this route is learned.
SPF No	This is the SPF calculation number that installed this route into the routing table.

The show ip route Command

The **show ip route** command is one of the most useful commands available for understanding and troubleshooting an IP network. Example 9-10 shows an example of a routing table in a multiarea network.

Example 9-1 *The show ip route Output*

```
Router#show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, * - candidate default
       U - per-user static route, o - ODR

Gateway of last resort is not set

    172.16.0.0/16 is variably subnetted, 3 subnets, 2 masks
O E2   172.16.20.128/29 [110/20] via 172.16.20.9, 00:00:29, Serial1
O IA   172.16.20.128/26 [110/74] via 172.16.20.9, 00:01:29, Serial1
C       172.16.20.8/29 is directly connected, Serial1
O E2  192.168.0.0/24 [110/20] via 172.16.20.9, 00:01:29, Serial1
```

Table 9-3 shows the link-state advertisement (LSA) codes used in the routing table, allowing you to analyze the flow of LSAs both within and between the areas.

Table 9-3 OSPF Routing Table Codes and Associated LSAs

LSA Type	Routing Table Entry	Description
1 Router Link	O	This is generated by the router, listing all the links to which it is connected, their status, and their cost. It is propagated within the area.
2 Network Link	O	This is generated by the designated router on a multiaccess network to the area.
3 or 4 Summary Link (between areas)	OIA	LSA Type 3 includes the networks or subnets within an area that might have been summarized and that are sent into the backbone and between ABRs. LSA Type 4 is information sent to the ASBR from the ABR. These routes are not sent into totally stubby areas.
5 Summary Link/ External Link (between autonomous systems)	OE1 or OE2	The routes in this LSA are external to the autonomous system. They can be configured to have one of two values. E1 will include the internal cost to the ASBR added to the external cost reported by the ASBR. E2 does not compute the internal cost—it just reports the external cost to the remote destination.

The show ip ospf virtual-links Command

The following command shows the configured virtual links that are in existence:

```
Router#show ip ospf virtual-links
```

Another command to use in conjunction with this is **show ip ospf neighbors** , which you learned about in Chapter 7.

Example 9-11 shows the output of the **show ip ospf virtual-links** command.

Example 9-1 The show ip ospf virtual-links Output

```
Router# show ip ospf virtual-links
Virtual Link to router 140.100.32.10 is up
Transit area 0.0.0.1, via interface Ethernet0, Cost of using 10
Transmit Delay is 1 sec, State DROTHER
Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
Hello due in 0:00:08
Adjacency State FULL
```

Table 9-4 explains the meaning of the important fields in the output of the **show ip ospf virtual-links** command.

Table 9-4 *Explanation of the show ip ospf virtual-links Command*

Field	Explanation
Virtual Link to router 140.100.32.10 is up	This shows the router ID of the other end of the virtual link, which is seen as a neighbor.
Transit area 0.0.0.1	This is the area through which the virtual link is tunneled: Area 0.0.0.1 or Area 1 in this case.
Via interface Ethernet0	This is the outgoing interface on the router that connects the virtual link to Area 0.
Cost of using 10	This is the cost of reaching the OSPF neighbor through the virtual link.
Transmit Delay is 1 sec	This is the delay of the link, how long it will take to transmit an LSA. This value must be less than the retransmit timer setting.
State DROTHER	This gives the state of the OSPF neighbor. The neighbor is a DROTHER, which is a router other than a designated router.
Hello 10	This gives the timed update interval for the Hello protocol, in seconds. The default is 10 seconds.
Dead 40	This tells how long the router will wait without hearing a Hello from the neighbor before it declares the neighbor dead. The default is 40 seconds.
Retransmit 5	The retransmit interval is the time in seconds that the router will wait without hearing an acknowledgment for the LSA that it has sent to its neighbor. The default is 5 seconds.
Hello due in 0:00:08	This shows the amount of time before the next Hello is expected from the neighbor.
Adjacency State FULL	This specifies the state of the neighbor adjacency. The two routers have fully synchronized their topological databases.

The show ip ospf database Command

The following command shows all the entries in the OSPF link-state database and the information taken from the LSAs that have been received. It can be tailored to show specific information from the database, such as the type 2 LSAs, otherwise known as the networks. To make the output specific to the LSA that you wish to examine, specify the appropriate parameter.

```
Router#show ip ospf database [router | network|summary|asbr-summary|nssa-external|external|database-summary]
```

Another command to use in conjunction with this is **show ip ospf neighbors** .

Example 9-12 shows the output of the **show ip ospf database** command.

Example 9-1 *The show ip ospf database Output*

```
Router#show ip ospf database

      OSPF Router with ID (172.16.20.130) (Process ID 100)

      Router Link States (Area 0)

Link ID      ADV Router    Age   Seq#  Checksum Link count
172.16.20.129 172.16.20.129 295   0x80000003 0x419B   1
172.16.20.130 172.16.20.130 298   0x80000002 0x3E9D   1

      Net Link States (Area 0)

Link ID      ADV Router    Age   Seq#    Checksum
172.16.20.130 172.16.20.130 298      0x80000001 0x19DB

      Summary Net Link States (Area 0)

Link ID      ADV Router    Age   Seq#    Checksum
172.16.20.8  172.16.20.129 291      0x80000001 0x7D1
```

Table 9-5 explains the meaning of the important fields in the output of the **show ip ospf database** command.

Table 9-5 *Explanation of the show ip ospf database Command*

Field	Explanation
Link ID 172.16.20.129	Router ID number
ADV Router 172.16.20.129	Advertising router ID
Age 295	Link state age
Seq# 0x80000002	Link state sequence number (detects old or duplicate LSAs)
Checksum 0x19DB	Checksum of the complete contents of the LSA
Link count 1	Number of interfaces detected for router

Troubleshooting a Multiarea OSPF Network

Troubleshooting OSPF across multiple areas is obviously more complicated than troubleshooting the configuration of a single area. Follow the basic tenets of troubleshooting to aid in unraveling problems and finding solutions in a timely manner:

- Maintain clear topology maps of the network.
- Keep current copies of all router configurations.
- Document changes made to the network.

Following these guidelines is imperative in a complex network to prevent it from becoming a labyrinth from which you cannot escape. The following sections cover the command **log-adjacency-changes** and the most useful **debug** commands. One of the most common problems that you might experience is the inability of neighbors to form adjacencies. You will also learn about the most common problems that occur in forming adjacencies, the reasons for them, and how to prevent them from occurring.

The log-adjacency-changes Command

A command that has a similar function to the **debug** command but does not require as many network resources is the **log-adjacency-changes** command. The **debug** command provides so much information that it is possible to overrun the logging buffers of the router as well as your own buffers. The **log-adjacency-changes** command offers less detail, which is often easier to assimilate. Whenever a change in the state of an adjacency occurs, a message is sent to syslog. This means that immediate updates are sent to the administrator without a massive drain on resources.

```
Router(config-router)#log-adjacency-changes
```

Example 9-13 shows the output of the **log-adjacency-changes** command. In this example, the OSPF process is reconfigured to log adjacency changes. The routing process for OSPF is started and the **network** command allocates all subnets to Area 0. Whether this is a reconfiguration of the area assignments or OSPF has never been run on this router before, the action is the same. The OSPF process reinitializes all interfaces, requiring the neighbors to be found and the adjacencies formed. With the adjacency changes logged, this example shows the interface Ethernet 0 finding the neighbors at 131.11.84.8 and 131.11.14.14 and all the steps to creating adjacencies being completed.

NOTE In Cisco IOS software release 12.1, the **ospflog-adjacency-changes** command became **log-adjacency-changes**.

Example 9-1 *OSPF log-adjacency-changes Output*

```

RouterA(config)#router ospf 1
RouterA(config-router)# log-adjacency-changes
RouterA(config-router)# network 0.0.0.0 255.255.255.255 area 0
RouterA(config-router)#end
RouterA#
10:30:15: %SYS-5-CONFIG_I: Configured from console by console
RouterA#
10:30:29: %OSPF-5-ADJCHG: Process 1, Nbr 131.11.14.14 on Ethernet0 from
DOWN to INIT, Received Hello
RouterA#
10:30:38: %OSPF-5-ADJCHG: Process 1, Nbr 131.11.84.8 on Ethernet0 from
DOWN to INIT, Received Hello
RouterA#
10:30:39: %OSPF-5-ADJCHG: Process 1, Nbr 131.11.14.14 on Ethernet0 from
INIT to 2WAY, 2-Way Received
RouterA#
10:30:48: %OSPF-5-ADJCHG: Process 1, Nbr 131.11.84.8 on Ethernet0 from
INIT to 2 WAY, 2-Way Received
RouterA#
10:30:54: %OSPF-5-ADJCHG: Process 1, Nbr 131.11.84.8 on Ethernet0 from
2WAY to EXSTART, AdjOK?
RouterA#
10:31:18: %OSPF-5-ADJCHG: Process 1, Nbr 131.11.84.8 on Ethernet0 from
EXSTART to EXCHANGE, Negotiation Done
10:31:18: %OSPF-5-ADJCHG: Process 1, Nbr 131.11.14.14 on Ethernet0 from
2WAY to EXSTART, AdjOK?
10:31:18: %OSPF-5-ADJCHG: Process 1, Nbr 131.11.14.14 on Ethernet0 from
EXSTART to EXCHANGE, Negotiation Done
10:31:18: %OSPF-5-ADJCHG: Process 1, Nbr 131.11.14.14 on Ethernet0 from
EXCHANGE to LOADING, Exchange Done
10:31:18: %OSPF-5-ADJCHG: Process 1, Nbr 131.11.14.14 on Ethernet0 from
LOADING to FULL, Loading Done
10:31:18: %OSPF-5-ADJCHG: Process 1, Nbr 131.11.84.8 on Ethernet0 from
EXCHANGE to LOADING, Exchange Done
RouterA#
10:31:18: %OSPF-5-ADJCHG: Process 1, Nbr 131.11.84.8 on Ethernet0 from
LOADING to FULL, Loading Done

```

Useful debug Commands

The **debug** commands are issued from the interactive console prompt, though remember that you need to be in privileged mode to issue the commands. Any fast-switched packets do not generate messages, so while debug is on, turn off fast switching and force all packets to be process switched for the few minutes that you are using the debug tool.

Use the following **debug** commands with caution. The resources required by the debug process can overwhelm your system.

The **debug ip packet** Command

The **debug ip packet** command is useful for analyzing the messages traveling between the local and remote hosts. IP debugging information includes packets received, generated, and forwarded.

The **debug ip ospf events** Command

Use the **debug ip ospf events** to display information on OSPF-related events, such as adjacencies, flooding information, designated router selection, and shortest path first (SPF) calculation.

Common Problems with Adjacencies

Many OSPF problems stem from adjacency problems that propagate throughout the network. Problems are often traced back to neighbor discrepancies. The following guidelines are helpful in these situations.

If a router configured for OSPF routing is not seeing an OSPF neighbor on an attached network, do the following:

- Make sure that both routers are configured with the same IP mask, MTU, interface Hello timer, OSPF Hello interval, and OSPF dead interval.
- Make sure that both neighbors are part of the same area and area type.
- Use the **debug** and **show** commands to trace the problem.

Foundation Summary

The “Foundation Summary” section of each chapter lists the most important facts from the chapter. Although this section does not list every fact from the chapter that will be on your exam, a well-prepared candidate should, at a minimum, know all the details in each “Foundation Summary” before going to take the exam.

Table 9-6 shows the codes used in the routing table.

Table 9-6 *OSPF Routing Table Codes and Associated LSAs*

LSA Type	Routing Table Entry	Description
1 Router Link	O	This is generated by the router, listing all the links to which it is connected, their status, and their cost. It is propagated within the area.
2 Network Link	O	This is generated by the designated router on a multiaccess network to the area.
3 or 4 Summary Link (between areas)	OIA	LSA Type 3 includes the networks or subnets within an area that may have been summarized and that are sent into the backbone and between ABRs. LSA Type 4 is information sent to the ASBR from the ABR. These routes are not sent into totally stubby areas.
5 Summary Link/External Link (between autonomous systems)	OE1 or OE2	The routes in this LSA are external to the autonomous system. They can be configured to have one of two values. E1 will include the internal cost to the ASBR added to the external cost reported by the ASBR. E2 does not compute the internal cost—it just reports the external cost to the remote destination.

Table 9-7 contains a list of the commands explained in this chapter. This list is not intended to teach the use of the commands, but to remind you of the options available.

Table 9-7 OSPF Command Summary

Command	Description
network <i>network-number wildcard-mask area area-id</i>	This command identifies the interfaces that are running OSPF and places them in the appropriate area.
summary-address <i>address mask [not-advertise][tag tag]</i>	This command consolidates routes into a summary route before injecting them into the external world. Remember that the mask is the subnet, not the wildcard mask. The option to not advertise will suppress routes that match the mask. The tag option is for use with redistribution. The summary-address command is also used to summarize external routes into OSPF, particularly during redistribution.
area <i>area-id range address mask</i>	This command is used to summarize the routes at the ABR before injecting them into another area.
no area <i>area-id range address mask</i>	This command disables the configured summarization at the ABR.
area <i>area-id stub</i>	This command turns the area into a stub area. This command must be configured on every router in the area.
area <i>area-id stub no-summary</i>	This command turns an area into a totally stubby area.
area <i>area-id default-cost cost</i>	This configures the cost for the default summary route used for a stub or totally stub area. The default is 1.
area <i>area-id virtual-link router-id</i>	This command creates a virtual link.
show ip ospf border-routers	This command lists the ABR and ASBR routers in the autonomous system.
show ip ospf virtual-links	This command shows the virtual links and the current parameters.
show ip ospf database	This command shows the OSPF database and the information taken from the LSAs that have been received.
debug ip ospf events	Displays information on OSPF-related events, such as adjacencies, flooding information, designated router selection, and SPF calculation. It verifies the value of the Hello timers on the interface.
debug ip ospf adjacencies	Displays information on the state of the adjacencies whenever there is a change of state.
log-adjacency-changes	Entered into the configuration file, beneath the routing process. This sends syslog messages in the event of a change in adjacencies. Fewer network resources are required in comparison to the debug commands.

Q&A

As mentioned in the introduction, “All About the CCNP, CCDP, and CCIP Certifications,” you have two choices for review questions. The questions that follow next give you a bigger challenge than the exam itself by using an open-ended question format. By reviewing now with this more difficult question format, you can exercise your memory better, and prove your conceptual and factual knowledge of this chapter. The answers to these questions are found in Appendix A.

For more practice with examlike question formats, including questions using a router simulator and multichoice questions, use the exam engine on the CD-ROM.

1. Which command in OSPF shows the network LSA information?
2. What command would you use to create a totally stubby area?
3. What is a virtual link, and what command would you use to create it?
4. Where would you issue the command to summarize IP subnets? State the command that is used.
5. How would you summarize external routes before injecting them into the OSPF domain?
6. When is a virtual link used?
7. Give the command for defining the cost of a default route propagated into an area.
8. Give an example of when it would be appropriate to define a default cost.
9. On which router is the area default cost defined?
10. Give the command to configure a stub area and state on which router it is configured.
11. What is the purpose of the **area range** command, and why is it configured on the ABR?
12. Give the commands to configure a router to place subnets 144.111.248.0 through to 144.111.255.0 in Area 1 and to put all other interfaces into Area 0.
13. Give the syntax to summarize the subnets 144.111.248.0 to 144.111.254.255 into another autonomous system.
14. Explain briefly the difference between the **area range** command and the **summary-address** command.
15. Explain the following syntax and what it will achieve: **area 1 stub no-summary** .
16. Why would you configure the routing process to log adjacency changes as opposed to turning on debug for the same trigger?
17. Give some of the common reasons that neighbors fail to form an adjacency.

18. When configuring a virtual link, which routers are configured?
19. What does the command **area 1 default-cost 15** achieve?
20. Explain what is placed in the parameters *area-id* and *router-id* for the command **area area-id virtual-link router-id**.

Scenarios

The following scenarios and questions are designed to draw together the content of the chapter and to exercise your understanding of the concepts. There is not necessarily a right answer. The thought process and practice in manipulating the concepts is the goal of this section. The answers to the scenario questions are found at the end of this chapter.

Scenario 9-1

Refer to Figure 9-9 and design the addressing scheme for the network. Then write the configuration for the central router.

Ensure that you include the following:

1. Address the network using the private network 10.0.0.0. Design the addressing scheme so that it allows for the summarization of addresses between areas. Show the summarization that you allocate, and explain your reasons for your choices.

Area 0 is using a prefix of 28 bits within the area.

Area 2 is using a prefix of 22 bits within the area.

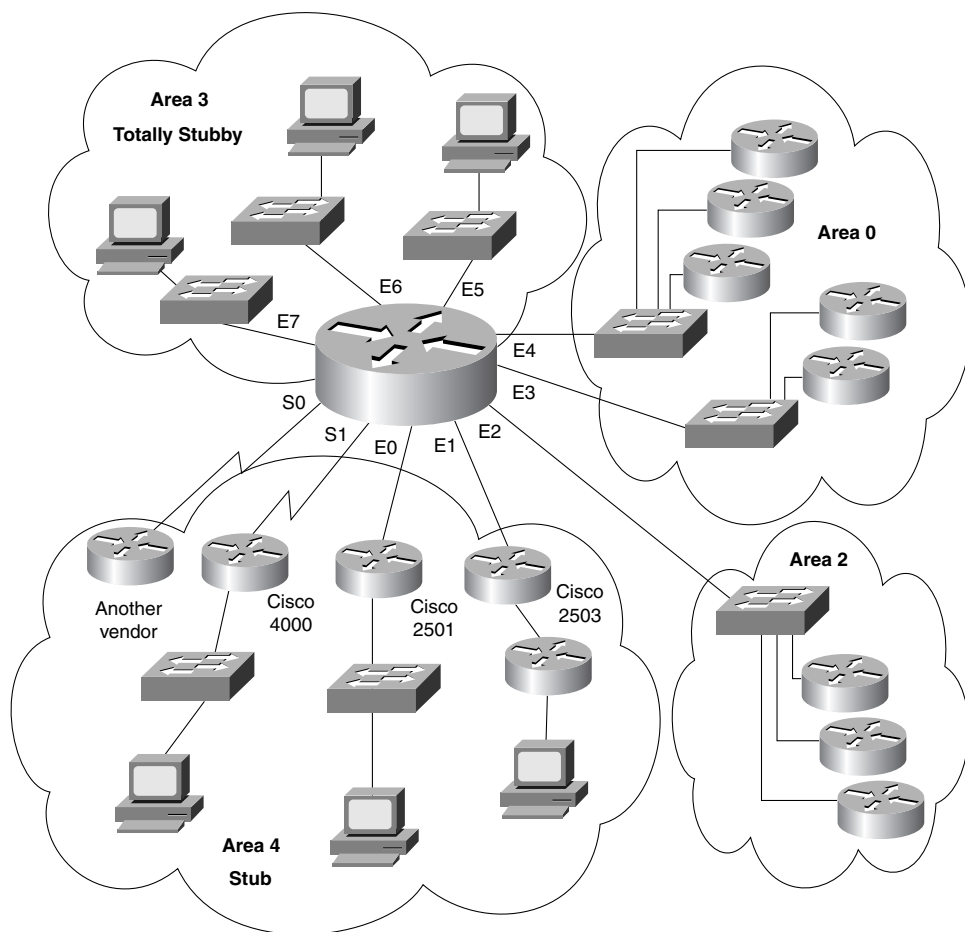
Area 3 is using a prefix of 24 bits within the area.

Area 4 is using a prefix of 30 bits for the serial connections. It is using a 28-bit prefix for the connections to the Ethernet routers. Do not include the subnets attached to the LANs in Area 4.

2. Issue the commands for the main router in Figure 9-9 to configure the following:

- The router ID
- The network commands to place the appropriate interfaces into the correct areas
- The configuration of the totally stubby area (Area 3)
- The configuration of the stub (Area 4)
- Summarization between areas
- The election of the central router as designated router, where appropriate

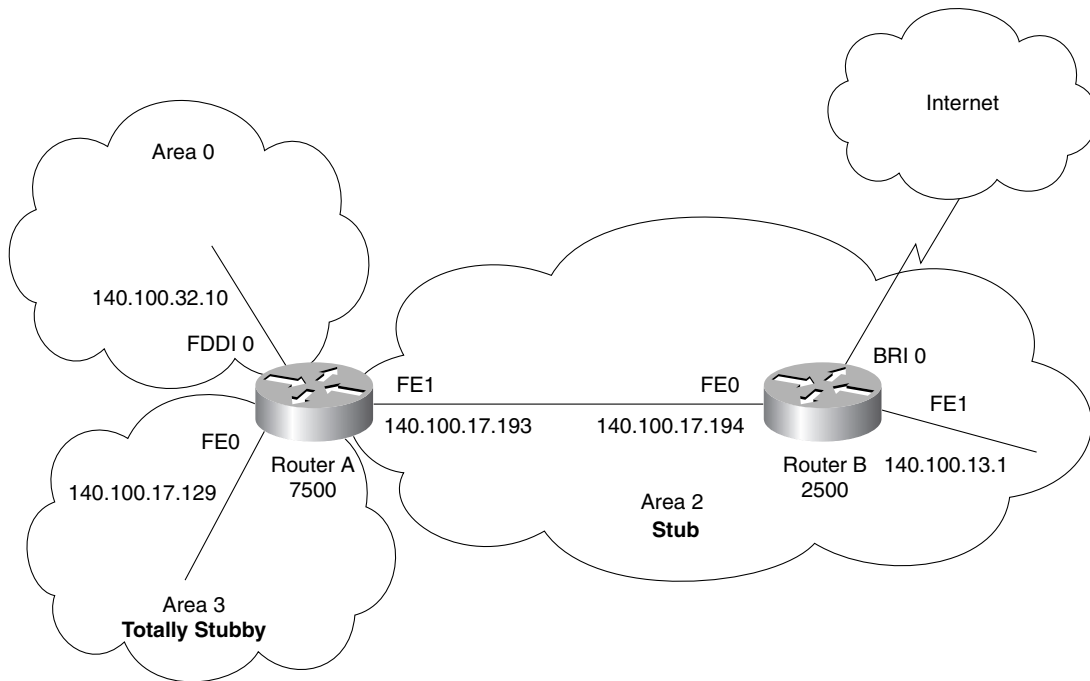
Figure 9-9 *The Diagram for Configuration Scenario 9-1*



Scenario 9-2

Use Figure 9-10 for this scenario.

The users of the network are complaining about the slowness of the network, particularly when trying to access the Internet. Examine the configuration in Example 9-14 in conjunction with Figure 9-10, and give reasons for any slowness or lack of connectivity that you can see on the network. Provide current configuration commands to correct any errors that you find.

Figure 9-10 *The Diagram for Configuration Scenario 9-2***Example 9-1** *Configuring OSPF Example for Scenario 9-2*

```

ROUTER A
Router(config)#router ospf 100
Router(config-router)#network 140.100.17.128 0.0.0.15 area 3
Router(config-router)#network 140.100.17.192 0.0.0.15 area 2
Router(config-router)#network 140.100.32.0 0.0.0.0 area 0
Router(config-router)#area 2 stub
Router(config-router)#area 3 stub no-summary
Router(config-router)#area 3 default-cost 15
!
Router(config-router)#interface FastEthernet0
Router(config-if)#ip address 140.100.17.129 255.255.255.240
Router(config-if)#no ip directed-broadcast
!
Router(config-if)#interface FastEthernet1
Router(config-if)#ip address 140.100.17.193 255.255.255.240
Router(config-if)#no ip directed-broadcast
!
Router(config-if)#interface Fddi0
Router(config-if)#ip address 140.100.32.10 255.255.255.240

```


Example 9-1 *Configuring OSPF Example for Scenario 9-2 (Continued)*

```

Router(config-if)#no ip directed-broadcast
Router(config-if)#no keepalive

Router(config-if)#exit
ROUTER B
Router(config)#router ospf 100
Router(config-router)# network 140.100.0.0 0.0.255.255 area 2
!
Router(config-router)#interface FastEthernet0
Router(config-if)#ip address 140.100.17.194 255.255.255.240
Router(config-if)#no ip directed-broadcast
Router(config-if)#ip ospf priority 100
!
Router(config-if)#interface FastEthernet1
Router(config-if)#ip address 140.100.13.1 255.255.255.240
Router(config-if)#no ip directed-broadcast
!
Router(config-if)#exit

```

Answer the questions that follow:

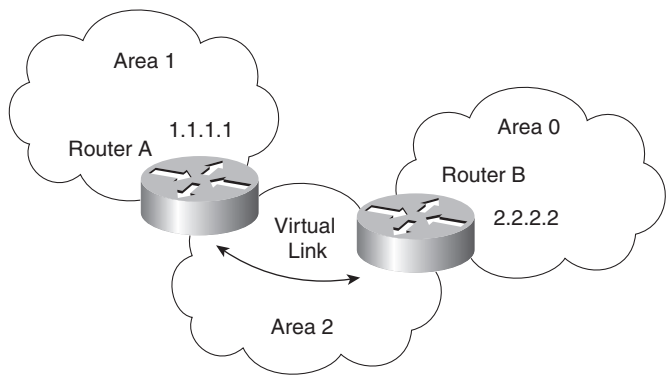
1. There are problems with Router B. There is inconsistency in the routing table, and the system is extremely slow. What commands would be used to identify the problem? In examining the diagram and configuration, what problems can you see?
2. Router A is having problems connecting to Area 0, which is causing problems in other areas because this router is used to connect to Area 0. What commands would be used to identify the problem? In examining the diagram and configuration, what problems can you see?
3. Issue the commands that would be used to correct the configuration problems that you see in the example configuration for Routers A and B.
4. When you issue the **show ip ospf** interface command, you notice that there is a discrepancy in the timers on the link between Routers A and B. The transmit timer on Router A is set to 5, and the retransmit timer is set to 1. What problems would this cause? What command would be used to change the timers, and what are the default settings?
5. There is an ISDN link into the Internet from Router B. The network manager has suggested that this link is the cause of some performance problems on the router. You have noticed that the interface is included in the OSPF **network** command. What might be the cause of the problem, and how could it be fixed?

Scenario 9-3

The network administrator is trying to solve a problem in the OSPF network and has determined that the creation of a virtual link is the answer. After studying the figures, answer the associated questions.

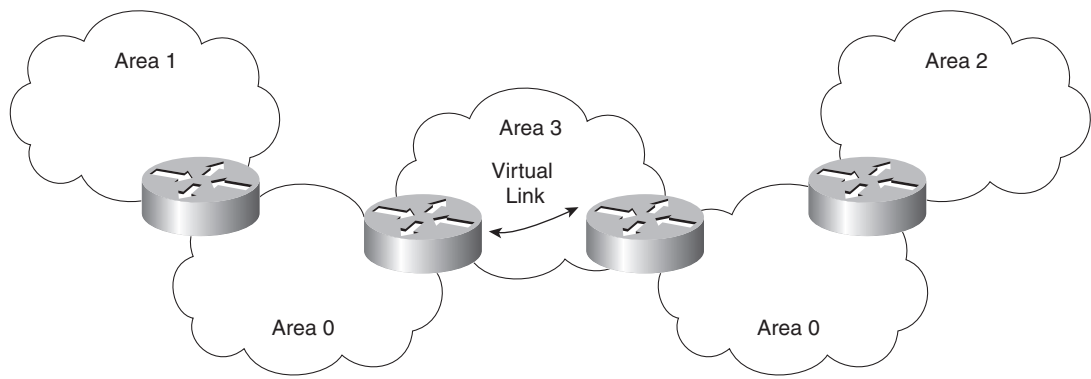
- 1. Explain the purpose of the virtual link in Figure 9-11.

Figure 9-11 Network Diagram 1 for Scenario 9-3



- 2. Is the configuration of the OSPF network shown in Figure 9-12 a valid configuration?

Figure 9-12 Network Diagram 2 for Scenario 9-3



- 3. Why would a company implement the design shown in Figure 9-12?

Scenario Answers

The answers provided in this section are not necessarily the only possible answers to the questions. The questions are designed to test your knowledge and to give practical exercise in certain key areas. This section is intended to test and exercise skills and concepts detailed in the body of this chapter.

If your answer is different, ask yourself whether it follows the tenets explained in the answers provided. Your answer is correct not if it matches the solution provided in the book, but rather if it has included the principles of design laid out in the chapter.

In this way, the testing provided in these scenarios is deeper: It examines not only your knowledge, but also your understanding and ability to apply that knowledge to problems.

If you do not get the correct answer, refer back to the text and review the subject tested. Be certain to also review your notes on the question to ensure that you understand the principles of the subject.

Scenario 9-1 Answers

Refer to Figure 9-9 and design the addressing scheme for the network. Then write the configuration for the central router.

Table 9-8 shows a possible addressing scheme using the criteria stated in Scenario 9-1. Taking the private address 10.0.0.0, there is a great deal of flexibility in the addressing scheme that can be devised. Remember, however, that careful filtering is required if the organization is to connect to the Internet.

The addressing scheme proposed here is broken out by area. It is not exhaustive in terms of designing an addressing policy down to the LAN level, as was shown in Chapter 3; instead, it deals with the principles of addressing and summarization.

Note in the allocation of addresses that this scenario deals with the allocation of subnets. The addresses displayed in this table are the address ranges of the available subnets, given the prefix length.

Table 9-8 Allocation of Addresses for Scenario 9-1

Area	Subnet/Prefix	Subnet Range	Reasons
0	10.0.0.0/28	1 0.0.0.16 to 10.0.255.224	The use of the 0s in the second octet is an easy reminder that you are in Area 0. Because Area 0 is a transit area, it will be small. The addressing within the area would be allocated the prefix of 28 bits, allowing the range of subnets shown. The number of hosts on each of these subnets would be 15.
2	10.2.0.0/22	10.2.0.0 to 10.2.252.0	Again, the private addressing of 10.0.0.0 as a Class A address is so large that full use can be made of the documentation advantages of the addressing scheme. The second octet allows Area 2 to be identified. The prefix of 22 bits is used within the area. This allows 1022 hosts on each network, which is good for further VLSM and VLANs in switched environments.
3	10.3.0.0/24	10.3.0.0 to 10.3.255.0	The second octet identifies the area. Within the area, a 24-bit prefix is used to address the LANs.
4	10.4.0.0/28 ... 10.4.100.0/30	Ethernet: 10.4.0.16 to 10.4.255.240 Serial: 10.4.100.0 to 10.4.100.250	The second octet identifies the area. Within the area, a 30-bit mask is used to identify the serial links on which only two addresses are needed. The subnet 10.4.100.0 was chosen as the subnet to sub-subnet for the serial links simply to ease troubleshooting; all serial links in the company in any area would be assigned the third octet of 100. A 28-bit mask was chosen for the Ethernet connections to allow the creation of many subnets.

Example 9-15 demonstrates a sample configuration for Scenario 9-1. The configuration file is for the central router.

Example 9-1 Sample Configuration of Scenario 9-1

```
Router(config)#router ospf 100
! Network Commands
Router(config-router)# network 10.0.0.0 0.0.255.255 area 0
Router(config-router)# network 10.2.0.0 0.0.255.255 area 2
Router(config-router)# network 10.3.0.0 0.0.255.255 area 3
Router(config-router)# network 10.4.0.0 0.0.255.255 area 4
! Totally Stubby Area
Router(config-router)# area 3 stub no-summary
```

Example 9-1 *Sample Configuration of Scenario 9-1 (Continued)*

```

! Stub Area
Router(config-router)# area 4 stub
! Summarization between areas
Router(config-router)# area 0 range 10.0.0.0 255.255.0.0
Router(config-router)# area 2 range 10.2.0.0 255.255.0.0
Router(config-router)# area 3 range 10.3.0.0 255.255.0.0
Router(config-router)# area 4 range 10.4.0.0 255.255.0.0
Router(config)# interface e0
Router(config-if)# ip address 10.4.0.33 255.255.255.240
Router(config)# interface e1
Router(config-if)# ip address 10.4.0.17 255.255.255.240
Router(config)# interface e2
Router(config-if)# ip address 10.2.4.1 255.255.252.0
!Election of Designated Router
Router(config-if)# ip ospf priority 64
Router(config)# interface e3
Router(config-if)# ip address 10.0.0.193 255.255.255.240
!Ensures Router not elected as Designated Router
Router(config-if)# ip ospf priority 0
Router(config)# interface e4
Router(config-if)# ip address 10.0.0.129 255.255.255.240
!Ensures Router not elected as Designated Router
Router(config-if)# ip ospf priority 0
Router(config)# interface e5
Router(config-if)# ip address 10.3.3.1 255.255.255.0
Router(config)# interface e6
Router(config-if)# ip address 10.3.2.1 255.255.255.0
Router(config)# interface e7
Router(config-if)# ip address 10.3.1.1 255.255.255.0
Router(config)# interface s0
Router(config-if)# ip address 10.4.0.9 255.255.255.252
Router(config)# interface s1
Router(config-if)# ip address 10.4.0.5 255.255.255.252
!Router ID set by configuring the Router ID with the loopback command
Router(config)# interface loopback 0
Router(config-if)# ip address 10.100.100.101 255.255.255.255

```

Scenario 9-2 Answers

1. *There are problems with Router B. There is inconsistency in the routing table, and the system is extremely slow. What commands would be used to identify the problem? In examining the diagram and configuration, what problems can you see?*

Router B has been configured to be the designated router for the LAN, which means that it is dealing with all the traffic on the LAN associated with the management of OSPF. Given that the system is a 2500, it is a poor choice for a designated router. A better choice would be Router A,

which is a larger system that connects directly to Area 0, making it a better choice from the standpoint of the network design. If Router B were a larger system than a 2500, there could be an argument for making it the designated router to elevate Router A, which would otherwise be functioning as the ABR as well as the designated router.

The router has not been configured as a stub, so the communication between Router A and Router B will be confused, preventing any communication between the two routers.

2. *Router A is having problems connecting to Area 0, which is causing problems in other areas because this router is used to connect to Area 0. What commands would be used to identify the problem? In examining the diagram and configuration, what problems can you see?*

Router A is configured incorrectly. The command that would show the problem would be either **show ip route**, **show ip protocols**, or **show ip ospf database**. The lack of LSA traffic would indicate a configuration problem. When examining the configuration, you would see that the mask on the configuration of the network command for Area 0 is wrong. The mask for 140.100.32.0 should be 0.0.0.15. Using the mask of 0.0.0.0 will place only interfaces with the IP address 140.100.32.0, which is a subnet address not an interface address. Therefore, there will be no communication of OSPF LSAs between the areas.

3. *Issue the commands that would be used to correct the configuration problems that you see in the example configuration for Routers A and B.*

The commands that would solve these problems are as follows:

On Router A:

```
router ospf 100
network 140.100.32.0 0.0.0.15 area 0
interface fastethernet 1
no ip ospf cost 10
ip ospf priority 100
```

On Router B:

```
router ospf 200
network 140.100.13.0 0.0.0.15 area 2
area 2 stub
interface FastEthernet0
no ip ospf priority 100
```

4. *When you issue the **show ip ospf interface** command, you notice that there is a discrepancy in the timers on the link between Routers A and B. The transmit timer on Router A is set to 5, and the retransmit timer is set to 1. What problems would this cause? What command would be used to change the timers, and what are the default settings?*

The default setting for the transmit timer is set to 1 second, and the retransmit timer is set to 5 seconds. The transmit timer determines the estimated number of seconds that it takes to send a LSA to a neighbor. The retransmit timer states the number of seconds to wait for an acknowledgment before retransmitting an LSA.

If the transmit timer is not smaller than the retransmit timer, the interface retransmits in the belief that the other side did not receive the LSA. This leads to excess traffic, confusion in the topology database, and the possibility of flapping links. To correct the settings, issue the following subinterface commands:

```
ip ospf retransmit-interval seconds
ip ospf transmit-delay seconds
```

5. *There is an ISDN link into the Internet from Router B. The network manager has suggested that this link is the cause of some performance problems on the router. You have noticed that the interface is included in the OSPF **network** command. What might be the cause of the problem, and how could it be fixed?*

If the ISDN interface is configured for dial-on-demand routing (DDR) and is also included in OSPF **network** commands, you might find that the link that the DDR process establishes will cause the routing updates to be propagated throughout the network. This causes additional CPU utilization on the routers and flooding of packets throughout the network. The solution is to ensure that the interface is not included in the **network** command to the OSPF process. A more important problem is that Router B is in a stub area and will not track external routes. Router B cannot connect to the Internet as an ASBR because it will not propagate the Type 5 LSAs. The BRI interface cannot partake in the OSPF network. Therefore, the network will not be slow; it will be down.

Scenario 9-3 Answers

1. *Explain the purpose of the virtual link in Figure 9-11.*

In this example, Area 1 does not have a direct physical connection into Area 0. A virtual link must be configured between Router A and Router B. Area 2 is to be used as a transit area, and Router B is the entry point into Area 0. This way, Router A and Area 1 will have a logical connection to the backbone.

2. *Is the configuration of the OSPF network shown in Figure 9-12 a valid configuration?*

Yes, the configuration is valid.

3. *Why would a company implement this design?*

OSPF allows for linking discontinuous parts of the backbone using a virtual link. In some cases, different Area 0s need to be linked together. This can occur, for example, if a company is trying to merge two separate OSPF networks into one network with a common Area 0. In other instances, virtual links are added for redundancy in case some router failure causes the backbone to be split in two. Whatever the reason may be, a virtual link can be configured between separate ABRs that touch Area 0 from each side and that have a common area between them.

Part III: IS-IS

Chapter 10 Fundamentals of the Integrated IS-IS Protocol

Chapter 11 Integrated IS-IS Protocol Operation

Chapter 12 Configuring Integrated IS-IS

Part III covers the following Cisco BSCI exam topics:

- Explain basic OSI terminology and network layer protocols used in OSI
- Identify similarities and differences between Integrated IS-IS and OSPF
- List the types of IS-IS routers and their role in IS-IS area design
- Describe the hierarchical structure of IS-IS areas
- Describe the concept of establishing adjacencies
- Given an addressing scheme and other laboratory parameters, identify the steps to configure Cisco routers for proper Integrated IS-IS operation
- Identify verification methods that ensure proper operation of Integrated IS-IS on Cisco routers
- Interpret the output of various **show** and **debug** commands to determine the cause of route selection errors and configuration problems



This chapter covers the following topics, which you need to understand to pass the CCNP/CCDP/CCIP BSCI exam:

- Introduction to Integrated IS-IS
- OSPF and IS-IS compared
- ISO addressing for Integrated IS-IS
- Integrated IS-IS hierarchical structure
- Basic principles of area routing
- Integrated IS-IS networks and interfaces
- Network layer protocols used in Integrated IS-IS

Fundamentals of the Integrated IS-IS Protocol

The topics in this chapter detail the routing protocol Integrated IS-IS. This chapter assumes knowledge of routing protocols and, in particular, link-state routing protocols. This chapter introduces Integrated IS-IS by explaining the protocol's terminology and fundamental concepts.

Because Integrated IS-IS is similar to the Open Shortest Path First (OSPF) protocol, you should read the chapters on OSPF and reinforce the fundamentals of link-state protocols. Differences between Integrated IS-IS and OSPF are clearly outlined within this chapter.

“Do I Know This Already?” Quiz

The purpose of the “Do I Know This Already?” quiz is to help you to decide what parts of this chapter to use. If you already intend to read the entire chapter, you do not necessarily need to answer these questions now.

The 21-question quiz, derived from the major sections in “Foundation Topics” portion of the chapter, helps you to determine how to spend your limited study time.

Table 10-1 outlines the major topics discussed in this chapter and the “Do I Know This Already?” quiz questions that correspond to those topics.

Table 10-1 “Do I Know This Already?” Foundation Topics Section-to-Question Mapping

Foundation Topics Section	Questions Covered in This Section
Introduction to Integrated IS-IS	1–3
OSPF and IS-IS Compared	4–6
ISO Addressing for Integrated IS-IS	7–9
Integrated IS-IS Hierarchical Structure	10–12
Basic Principles of Area Routing	13–15
Integrated IS-IS Networks and Interfaces	16–18
Network Layer Protocols Used in Integrated IS-IS	19–21

NOTE The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark this question wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

1. What does the initialism CLV stand for?
 - a. Complete/Length/Verification
 - b. Code/Length/Value
 - c. Current/Long/Vector
 - d. Code/Length/Vector
2. What do the initials CSNP represent?
 - a. Current System Node Packet
 - b. Complete Sequence Number Protocol
 - c. Code Sequence Number Protocol
 - d. Complete sequence number packet
3. What do the initials CLNS represent?
 - a. Connectionless Network Service
 - b. Connection Network Service
 - c. Code/Network/Service
 - d. Complete Network Sequence
4. Which of the following do IS-IS and OSPF have in common?
 - a. Dijkstra algorithm
 - b. Classless routing protocol
 - c. Link-state routing protocol
 - d. Summarization at the area boundary
5. OSPF uses the term autonomous system. Which of the following is the equivalent IS-IS term?
 - a. Area
 - b. Routeing Domain

- c. Initial domain part
 - d. Authority format identifier
6. OSPF assigns the network address to the interface. To what is the IS-IS address assigned?
- a. To the CLNS interface
 - b. To the IANA network address
 - c. To the interface
 - d. To the IS-IS process
7. ISO 10589 defines the ISO address as having three fields. What are these fields?
- a. IDP, DSP, SEL
 - b. Area, System ID, and SEL
 - c. IDP, DSP, System ID
 - d. AFI, IDP, HODSP
8. Which of the following is a valid NET address?
- a. 47.0005.aa00.0301.16cd.00
 - b. 47.0005.aa00.0301.16cd.01
 - c. 47.0005.aa00.0301.16cd.ff
 - d. 47.0005.aa00.19g6.3309.00
9. In the rules for ISO addressing, the System ID for a Level 2 router must be unique at which of the following levels?
- a. The interface on the router
 - b. The area
 - c. The routing domain
 - d. Everywhere
10. A Level 1 router is described as which of the following?
- a. An interarea router
 - b. An HODSP router
 - c. An internal router
 - d. An intra-area router

11. How many IS-IS processes run on a Level 1-2 router?
 - a. One
 - b. Two
 - c. One for each area
 - d. Two for each area
12. A Level 2 router is similar to which type of OSPF router?
 - a. ASBR
 - b. ABR
 - c. Stub
 - d. Backbone internal router
13. Where are IS-IS areas defined?
 - a. On the link
 - b. On the interface
 - c. At the process
 - d. A separate process is required for each routing table
14. Which of the following statements is true about Level 2 routers?
 - a. Level 2 routers can send updates between routing areas.
 - b. They must be placed contiguously.
 - c. They are similar to OSPF stub routers.
 - d. This is the Cisco default configuration.
15. Once a datagram is accepted into the router, which of the following fields are stripped?
 - a. CRC
 - b. System ID
 - c. SEL
 - d. Protocol type
16. For an adjacency to be formed and maintained, both interfaces must agree on which of the following?
 - a. If the routers are both Level 1, they must be in the same area.
 - b. The system ID must be unique to each router.

- c. Each router needs to be configured as the same level of routing—that is, either Level 1 or Level 2.
 - d. The Hello timers.
- 17.** How does the pseudonode represent all the routers on the LAN?
- a. Each router on the multiaccess LAN simulates an interface on the pseudonode.
 - b. The pseudonode represents the physical link to which the routers are connected.
 - c. The pseudonode is the virtual link between two discontinuous areas.
 - d. Each link on every router on the LAN is summarized to form a pseudonode for the LAN.
- 18.** What is used to elect the DIS automatically?
- a. Highest SNPA address.
 - b. Priority defined at the interface.
 - c. The DIS must be manually configured.
 - d. The highest loopback address.
- 19.** Which of the following packets are used by IS-IS?
- a. Hellos
 - b. LSPs
 - c. TVLs
 - d. SNPs
- 20.** Which of the following are valid hello packet types for IS-IS?
- a. Hello Level 1 LAN
 - b. Hello Level 2 LAN
 - c. Hello Level 1-2 LAN
 - d. Hello Point to Point
- 21.** IS-IS uses the Hello protocol to create and maintain adjacencies and neighbor relations. Which of the following are Hello packets used in IS-IS?
- a. A generic Hello
 - b. Point-to-point Level 1, point-to-point Level 2, LAN Level 1, LAN Level 2
 - c. Point-to-point, LAN Level 1, LAN Level 2
 - d. Point-to-point Level 1, point-to-point Level 2, LAN

The answers to this quiz are found in Appendix A, “Answers to Chapter ‘Do I Know This Already?’ Quizzes and Q&A Sections.” The suggested choices for your next step are as follows:

- **12 or less overall score** —Read the entire chapter. This includes the “Foundation Topics,” “Foundation Summary,” and “Q&A” sections at the end of the chapter.
- **13–18 overall score** —Begin with the “Foundation Summary” section and then go to the “Q&A” section at the end of the chapter. If you have trouble with these questions, read the appropriate sections in “Foundation Topics.”
- **19 or more overall score** —If you want more review on these topics, skip to the “Foundation Summary” section and go to the “Q&A” section at the end of the chapter. Otherwise, move to the next chapter.

Foundation Topics

Introduction to Integrated IS-IS

IS-IS is an Interior Gateway Protocol (IGP) developed in the 1980s by Digital Equipment and submitted to the International Organization for Standardization (ISO) as the routing protocol for Open System Interconnection (OSI). The creation of IS-IS was an attempt to produce a standard protocol suite that could allow internetworks to scale.

The development of IS-IS was motivated by the need for the following:

- A nonproprietary protocol
- A large addressing scheme
- A hierarchical addressing scheme
- A protocol that was efficient, allowing for fast, accurate convergence and low network overhead

The United States mandated that every system operated by the government had to be capable of running the OSI architecture (as an initiative called the Government Open Systems Interconnections Profile [GOSIP]). By forcing every government system to understand OSI, officials hoped that the protocols would become the practical standard and the academic solution to diverse proprietary implementations for computer networking. In the end, however, the initiative failed. The Internet, built on TCP/IP, prevailed as the practical alternative to an international standard. However, IS-IS has always been used, although not as extensively as first hoped. Large Internet service providers (ISPs) have been using IS-IS since its inception in the 1980s, and recently IS-IS has begun to emerge in other markets. This new interest is for a variety of reasons, including the fact that IS-IS is a standard that provides protocol independence, it has the capability to scale, and it has the capacity to define type of service (ToS) routing, though currently this is not a supported feature in the Cisco IOS. ToS routing allows traffic engineering, which requires very complex routing decisions to be programmed into the protocol. Therefore, as of late, IS-IS has been taken from the shelf, dusted off, and put into use.

IS-IS Terminology

The terminology used by IS-IS might appear cumbersome and unfriendly. Although the jargon might be unfamiliar, most of the concepts are no different from other routing protocols. Table 10-2 explains briefly some of the commonly used IS-IS terms.

Table 10-2 *IS-IS Terms*

Term	Definition
Adjacency	Local routing information that shows the reachability of a directly connected ES or IS. A separate adjacency is created for each neighbor on a circuit, and for each level of routing (that is, Level 1 and Level 2) on a broadcast circuit.
Administrative Domain	A group of routers that share the same routing protocol within one organization.
Area	A subdomain within an Administrative Domain. Routers in an area maintain detailed routing information about the area's internal composition. The routers also maintain routing information that allows them to reach other areas. The area address is contained in the NET and NSAP address.
Circuit	The local routing information for a single subnet point of attachment (SNPA).
Code/Length/Value (CLV)	These are the variable-length fields in a PDU. The Code field specifies the information in the Content field as a number. The Length field states the size of the Value field. The Value field contains the information itself.
Complete sequence number packet (CSNP)	CSNPs describe every link in the link-state database. CSNPs are sent on point-to-point links when the link comes up to synchronize the link-state databases. The designated router (DR), or designated intermediate system (DIS), on a multicast network sends out CSNPs every 10 seconds.
Connectionless Network Protocol (CLNP)	This is the ISO protocol used to carry data and error indications at the network layer. CLNP is similar to IP and has no facilities to detect errors in data transmission. It relies on the transport layer to provide guaranteed data delivery.
Connectionless Network Service (CLNS)	CLNS uses a datagram transfer service and does not require a circuit to be established before data is transmitted.

Table 10-2 *IS-IS Terms (Continued)*

Term	Definition
Connectionless Network Service (CLNS) (continued)	Whereas CLNP defines the actual protocol, CLNS describes a service provided up to the transport layer. Being a connectionless service, CLNP provides a “best effort” delivery of data; therefore, there is no guarantee that data will not be lost, corrupted, misordered, or duplicated. If you require guaranteed delivery, the transport layer or application layer needs to provide the service that will correct the problems when they arise.
Designated intermediate system (DIS)	The router (IS) on a LAN that is designated to perform additional duties. In particular, the DIS generates link-state PDUs on behalf of the LAN by treating the LAN as a pseudonode.
Dual IS-IS	IS-IS that supports both OSI and IP routing information. Areas within the autonomous system can run either OSI or IP or both. However, the configuration chosen must be consistent within the entire area.
End system (ES)	The end node or host, which has limited routing capabilities. The ES has the OSI or IP Layer 3 protocol running and can receive and send data.
End System-to-Intermediate System (ES-IS)	The protocol by which the OSI ES and the IS communicate to dynamically learn Layer 2 adjacencies.
Hello	Hello packets are used to discover and maintain adjacencies.
Host address	This is a subset of the NET address, which includes both the domain, area, and system ID.
Integrated IS-IS	Another term for Dual IS-IS. Indicates IS-IS can be used to support routing for two Layer 3 protocols (IP and CLNP) in the same network simultaneously.
Intermediate system (IS)	A router. The IS is a device capable of directing traffic to remote destinations.
Intermediate System-to-Intermediate System (IS-IS)	The OSI routing protocol that learns the location of the networks within the autonomous system so that data can be forwarded to the remote hosts.

(continues)

Table 10-2 *IS-IS Terms (Continued)*

Term	Definition
IS-IS domain	A group of routers running the IS-IS protocols for exchanging routing information.
Level 1 (L1)	These routers are internal to the area, which means that they receive routing information for their area only and have no knowledge of the other areas' networks. To reach other areas, Level 1 routers maintain a default route to the nearest Level 2 router.
Level 1-2 (L1-2)	A router that connects areas. This router connects a Level 1 area to the Level 2 backbone. It will have a Level 1 routing table to route to ES and IS in its own area by system ID. It will maintain a Level 2 prefix table to route to other areas.
Level 2 (L2)	These routers are connected only to the backbone and provide transit traffic between areas.
Link	A physical connection to a neighbor. This link is then transmitted to all the other routers in the area via the LSP.
Link-state packet (LSP)	A packet that describes a router's links. There are separate LSPs for Level 1 and Level 2 updates.
Neighbor	A router on the same link with which an adjacency is formed and routing information is then exchanged.
Network entity title (NET)	Part of the OSI address. The NET describes both the area and system ID of a system in the IS-IS network but excludes the NSEL, which defines the NSAP address of the system.
Network protocol data unit (NPDU)	See protocol data unit (PDU).
Network selector (NSEL)	Sometimes referred to as the SEL field. This field describes the service at the network layer by which the packet is to be sent. NSEL is similar to the Protocol field in IP.
Network service access point (NSAP)	Describes a service at the network layer to which the packet is to be directed. The NSAP is the NET address with the SEL field set to a value other than 0x00.

Table 10-2 *IS-IS Terms (Continued)*

Term	Definition
Overload (OL) bit	The OL is set on an LSP if the router cannot store the entire link-state database. When other routers receive LSPs with this bit set, they will not send the router any transit traffic for fear that its routing table is incomplete. If the router is making decisions using incomplete data, its decisions may result in suboptimal paths or even routing loops. Traffic destined for the router can still be sent to the directly connected interfaces of a router transmitting the OL bit in its LSPs.
Partial sequence number packet (PSNP)	PSNPs are sent on point-to-point links to acknowledge explicitly each LSP the router receives. A router on a broadcast subnetwork sends a PSNP requesting the LSPs it needs to synchronize its link-state database.
Protocol data unit (PDU)	A unit of data passed from one layer of the OSI model to the same level of the OSI model on another node. Each layer prefixes the PDU to indicate the sending OSI layer so that the network layer sends NPDUs and the data-link layer sends DLPDUs.
Pseudonode	The LAN identifier for a broadcast subnetwork. The pseudonode makes the broadcast medium appear as a virtual router and the routers appear as connected interfaces. The routers maintain adjacencies to the pseudonode, which are managed by the DIS, instead of to all other routers on the medium (thus reducing memory, CPU, and bandwidth resources).
Routeing Domain	Routeing Domain is the same as the Administrative Domain. It defines the boundaries of a network of interconnected routers operated and managed by the same administrative group. The spelling of Routeing Domain is not a typographic error but the British spelling adopted by the ISO committee.
Sequence number PDU (SNP)	SNPs are used to acknowledge the receipt of LSPs and to synchronize link-state databases.

(continues)

Table 10-2 *IS-IS Terms (Continued)*

Term	Definition
Subnetwork	The data-link layer.
Subnetwork dependent layer	Interfaces with the data-link layer and hides the different kinds of data-link layers from the network layer. This sublayer transmits and receives PDUs from the subnetwork, translates DLPDUs into NPDUs, and hands them to the appropriate OSI process. The subnetwork dependent layer is also responsible for creating and maintaining adjacencies through the exchange of IS-IS Hello PDUs.
Subnetwork independent layer	Interfaces with the transport layer and provides it with network services. It describes how CLNS creates and maintains knowledge of the network by exchanging and processing routing information so that data can be transmitted efficiently to remote destination hosts and handed to the transport layer.
Subnetwork point of attachment (SNPA)	The data-link layer offers two services: the physical connection to the medium and the services offered to the physical layer and network layer. The SNPA refers to these services. The SNPA address is the physical address (for example, the MAC address on a LAN).
Type/Length/Value (TLV)	TLV is the same as a CLV, but some literature refers to the variable-length fields as TLV in accordance with the IP terminology.

OSPF and IS-IS Compared

Integrated IS-IS and OSPF share a common heritage. The following sections compare the similarities and differences between the two routing protocols.

Similarities Between Integrated IS-IS and OSPF

They are both link-state protocols and are based on the Dijkstra algorithm of Shortest Path First (SPF). In addition, they both have a two-level hierarchy. OSPF tends to be deployed mostly as an enterprise solution, whereas Integrated IS-IS is used for IP routing in some ISP networks.

Table 10-3 outlines other similarities between Integrated IS-IS and OSPF.

Table 10-3 *Integrated IS-IS and OSPF Similarities and Terminology Comparison*

IS-IS Terminology	OSPF Terminology
Area	Stub area
Area ID	Area ID
Backbone area	Backbone area
DIS (designated intermediate system)	Designated router
Domain	Network
ES (end system)	Host
ES-IS (the address resolution feature of ES-IS)	ARP (Address Resolution Protocol)
IS (intermediate system)	Router
ISO Routing Domain	Autonomous system
Level 1	Internal nonbackbone stub area
Level 1-2	Area border router (ABR)
Level 2	Backbone router
LSP (link-state packet)	LSA (link-state advertisement)
CSNP and PSNP (complete and partial sequence number PDUs)	Link-state acknowledgement packet
PDU (protocol data unit)	Packet
NET (network entity title)	IP destination address (subnet and host), used in a similar way to router ID
NSAP (network service access point)	IP destination address + IP protocol number

(continues)

Table 10-3 Integrated IS-IS and OSPF Similarities and Terminology Comparison (Continued)

IS-IS Terminology	OSPF Terminology
Routing technology = link state • Classless routing protocol • Address summarization between areas • Uses a link-state database • Acknowledges LSPs • Shortest path is computed using Dijkstra’s SPF algorithm • Hellos create and maintain adjacencies • Hellos and holdtime may be configured	Routing technology = link state • Classless routing protocol • Address summarization between areas • Uses a link-state database • Acknowledges LSAs • Shortest path is computed using Dijkstra’s SPF algorithm • Hellos create and maintain adjacencies • Hellos and holdtime can be configured
Subnet = data link	Subnet = IP network
SNPA (subnetwork point of attachment)	Layer 2 address; for example, the MAC address (Media Access Control) or DLCI (data-link connection identifier) in Frame Relay
System ID	Address of the host within the network, sometimes used as the router ID determining priorities
Virtual link (defined but not supported)	Virtual link

Differences Between OSPF and Integrated IS-IS

Although OSPF and Integrated IS-IS share the same common goals and use the same link-state technology to achieve those goals, the methods they use differ slightly. For example, the protocols differ in how the area address is assigned. In IS-IS, the area and host address are assigned to the entire router, whereas in OSPF, the address is assigned at the interface level. An IS-IS router is therefore in one area, while in OSPF, a router can inhabit many areas, though there is an option for multiarea IS-IS that is used primarily during area migrations and transition.

This means that all Level 1 routers have to be within the same area, with a Level 1-2 router connecting them to another area. However, the Level 1-2 router needs to be in the same area as the Level 1 router with which it communicates. The Level 1-2 router can see the rest of the autonomous system and offers itself as the default route to the Level 1 area. This is similar to the OSPF stub areas. The Level 2 router sends Level 2 updates in the other area, or prefix routes, just like the ABR in OSPF.

The role of the DR is subtly different. The DIS in IS-IS exists for both Level 1 and Level 2 on multiaccess media, but there is no backup designated router (BDR). Also, in OSPF the DR is elected for life; in IS-IS, however, if another router comes on line with a higher priority, the existing DIS is

deposited. Fewer adjacencies are formed in OSPF because the routers form adjacencies only with the DR and the BDR. In IS-IS, every router makes an adjacency with every other router on the medium. However, IS-IS LSPs are sent out only by the DIS on behalf of the pseudonode.

A major difference is the encapsulation of the two protocols. IS-IS has protocol independence because it runs directly on top of the data-link layer. Fragmentation is the responsibility of the IS-IS process, and this allows for a streamlined protocol. More importantly, fragmentation makes evolution of the protocol as needed very simple, because it is not dependent and therefore is not limited by any third protocol. OSPF is encapsulated into IP and is limited by the capabilities of that protocol.

The way that the LSPs are handled is also slightly different and influence, to an extent, the design of a network running either protocol. Unrecognized LSPs are ignored and flooded in IS-IS, for example; OSPF ignores and drops unrecognized LSAs.

Table 10-4 lists the key differences.

Table 10-4 *Integrated IS-IS and OSPF Technical Differences*

Technology	Integrated IS-IS	OSPF
Areas	Boundaries defined on the link. A router (IS) can be in one area, though there is an option for multiarea IS-IS that is used primarily during area migrations and transition	Area boundaries are defined on the router. Interfaces can be in different areas. A router might be in many areas. The Integrated IS-IS Level 1 area is similar to an OSPF stub area.
Designated router (DR)	If a new IS becomes active higher priority, it becomes the new DIS, which results in a flood of LSPs. If the priority of the new router is the same and the MAC address is higher, it takes over the DIS functions. Adjacencies are created with all ISs on the broadcast media. Each IS sends a multicast LSP to all ISs on the media. The LSP is unacknowledged.	A same or higher priority does not dislodge the existing DR. Adjacencies on broadcast media are formed with the DR and BDR only. All LSAs are acknowledged.

(continues)

Table 10-4 Integrated IS-IS and OSPF Technical Differences (Continued)

Technology	Integrated IS-IS	OSPF
Encapsulation	Integrated IS-IS runs on top of the data-link layer (Layer 2). Integrated IS-IS is a network layer protocol with its own Layer 3 packet. Fragmentation is the responsibility of Integrated IS-IS.	OSPF is an IP application. Has an OSPF header and travels inside an IP packet. Fragmentation is the responsibility of IP.
LAN flooding	All ISs maintain adjacencies with all other ISs on a broadcast network. DIS sends CSNP to all ISs. Periodic CSNPs ensure the databases are synchronized.	Multicast updates and Hellos sent to DRs. Unicast acknowledgment sent from the DRs.
LSAs	Two types of LSP. LSPs are CLV encoded. Unrecognized LSPs are ignored and flooded. LSPs are always flooded across all media by the originating IS.	Seven types of LSA. Unrecognized LSAs are not flooded. Many small LSAs for summary and external updates. LSA updates generated by each router.

ISO Addressing for Integrated IS-IS

Because IS-IS is the product of a committee, it has the feel of an academic solution that is intended to resolve every eventuality. Its addressing scheme thinks not just locally, but globally. The large address space is one characteristic that is responsible for IS-IS’s new popularity.

When Integrated IS-IS is routing IP traffic, the routing information is carried in the IS-IS updates; therefore, the participating routers need an ISO address. The ISO address comes in two forms: the NSAP and the NET address (depending on the type of device being addressed). The addressing scheme of IS-IS is long and complex, but clear rules define the address space, allowing the IS-IS protocol to find the routers to forward the data traffic to the end destination.

The IS-IS address is a variable-length address from 8 to 20 octets in length, much larger than the 4 octets in the TCP/IP address. ISO 10589 defines three parts to the address—Area, ID, and SEL—as shown in Figure 10-1.

Figure 10-1 The Three Sections of an ISO Address

AREA	ID	SEL

IS-IS addressing can become complicated because these three fields are subdivided to allow greater granularity in routing. If you are confused, however, just remember that it all condenses to the three elements of the address, which perform the following functions:

- **Area** — The Area field is used to route between areas using Level 2 routing.
- **ID** — The ID field is used to route to a host or router within the area using Level 1 routing.
- **SEL** — The SEL field is used to route to an entity within the host or ES.

Although the three parts to the address describe how to get to the area, how to find the host, and how to find the application within the host, finding the destination host uses only the first two parts; the last part of the address is used after the end host has received the packet. Therefore, Integrated IS-IS has two levels of hierarchy: basically, how to get to the area and then to the host.

The first part of the address, routing to the area, might require many decisions to be made. These decisions might involve determining the country and then the organization and many suborganizations. So although the address has only two levels of hierarchy, several levels of hierarchy can be hidden within the first part of the address, which states how to get to the autonomous system and forms the external portion of the area.

Figure 10-2 shows the division between external and internal routing — the Initial Domain Part (IDP) and the Domain Specific Part (DSP) — and how this division is broken down to reveal greater granularity for routing. In the figure, the three original elements have been placed underneath as reference.

Figure 10-2 An ISO Address

IDP		DSP		
AFI (1 octet)	IDI	High Order DSP	System ID (1-8 octets)	NSEL (1 octet)
AREA			ID	SEL

The following list explains the parts of an ISO address and their relationship:

- **IDP** — External routing. The IDP is used to route to the domain, or autonomous system. The IDP is given by the ISO and identifies the organization responsible for assigning the format for the rest of the address by defining the DSP structure. The IDP comprises the following two parts:
 - **Authority and format identifier (AFI)** — AFI is the first octet of the address. It is defined as one octet. It states the authority in charge of issuing addresses and the

format of the rest of the address, within the constraints of IDI and DSP, to describe Area, ID, and SEL.

- **Initial domain identifier (IDI)** — IDI is a suborganization to the AFI; as an analogy, the U.S. government is the AFI, and the General Services Agency (GSA) is the IDI, otherwise known as GOSIP.
- **DSP** — DSP is used to route within the autonomous system. The authority specified by the IDP may further subdivide the DSP and may assign subauthorities responsible for parts of the DSP. The DSP can be broken down into the following fields:
 - **High Order DSP** — High Order DSP is typically the area within the autonomous system.
 - **System ID** — System ID can have a value between one to eight octets. It must have the same length throughout all systems in the autonomous system. Cisco uses six octets, which is a common solution because it allows the use of the MAC address to autoconfigure the system.
 - **NSEL** — NSEL is one byte and identifies the particular service at the network layer to which to hand the packet.

The fact that the address can take so many forms can cause confusion; remember, however, that there are only two layers of hierarchy. By providing such flexibility in the address space, the ISO has ensured a decentralized address allocation and management, in addition to the ability to scale the network.

As with TCP/IP, the addressing scheme within an autonomous system can be the result of the creative genius of the network administrator or can be obtained from the AFI, an authorized ISO body such as ANSI or GOSIP.

The next sections discuss ISO addresses for Integrated IS-IS, including examples of NETs and NSAP, in addition to the rules for IS-IS addressing.

NETs and NSAP

NETs and NSAPs are ISO addresses. The differences between the NET and NSAP addresses are subtle. The NET address is the address of the host, where the value in the NSEL field is set to 0x00. Therefore, there is no upper-layer protocol identified within the host. With no application identified within the end host, the packet can be routed to the destination, but it cannot be handed off to a process after it has been delivered. However, routers do not have upper-layer protocols to identify because they are transitory ISs. Therefore, the NSAP of the router is referred to as a NET because the NSEL field is set to 00.

The NSAP is the full ISO address. It not only defines the area and destination host within the area, but also specifies where to send the incoming packet after it has reached the host. The NSEL field

at the end of the ISO address specifies the upper-layer protocol and is similar to the Protocol field of the IP header.

Rules of ISO Addressing

The following list indicates a few rules that clarify ISO addressing:

- The ISO address is assigned to the system, not to the interface.
- Typically, the router has one NET address. The limit of three NETs is in a conventional IS-IS implementation; the limit is three NETs per area in a multiarea Integrated IS-IS implementation. Multiple addresses are used during transitions.
- If multiple NETs are configured on the same router, they must all have the same system ID.
- The area address must be the same for all routers in the same area.
- All Level 2 routers must have a system ID that is unique for the entire domain.
- All Level 1 routers must have a system ID that is unique for the entire area.
- The system ID must be the same length for all ISs and ESs within a routing domain.

Example of a NET Address

The following are examples of NET addresses. The first two addresses are used for routing within the autonomous system and, therefore, the IDI portion of the address has not been defined.

- A simple OSI NET address, using the host MAC address as the system ID:

47.0005.aa00.0301.16cd.00

← To the Domain →		← Within the Domain →		
AFI	IDI	HO-DSP	System ID	SEL
47.		0005.	aa00.0301.16cd.	00
Area			System ID	SEL

- A simple OSI NET address, using the host's loopback IP address of 144.132.16.19 as the system ID:

47.0001.1441.3201.6019.00

← To the Domain →		← Within the Domain →		
AFI	IDI	HO-DSP	System ID	SEL
47.		0001.	0001.1441.3201.6019.	00
Area			System ID	SEL

- A GOSIP version 2 address, showing the external routing information:

47. 0005.80ff.f800.0000. 0001.0000.0c00.1234.00

← To the Domain		→← Within the Domain →		
AFI	IDI	HO-DSP	System ID	SEL
47.	0005.80ff.f800.0000.	0001.	0000.0c00.1234.	00
Area			System ID	SEL

Integrated IS-IS Hierarchical Structure

As the addressing structure shows, there are potentially many levels of hierarchy within the addressing scheme, even though the protocol has only two layers of hierarchy. To accommodate the two levels of hierarchy, two types of router are defined: a Level 1 router that deals with the first level of routing, finding the end destination within the area, and a Level 2 router that finds the area within which the end destination resides. Both of these routers are combined in the Level 1-2 router, which runs both the Level 1 and the Level 2 processes and could be viewed as a third type of router.

The Level 1 Router

The Level 1 router locates the destination host within the area and, as such, is known as the intra-area router. It is similar to a stub router in OSPF. Like a stub router in OSPF, its knowledge of the network is limited to the area, using a default route to the nearest Level 2 router as the means of routing traffic external to the area.

Every Level 1 router has a link-state database containing all the routing information for the area. The nature of the link-state protocol means that each database is identical. Because the database is limited to routes within the area, there is a requirement that neighbors must be in the same area so that the routers can communicate.

The Level 2 Router

To route traffic between areas, a Level 2 router is needed. Routing between areas is referred to as *interarea routing*. This router is similar to a backbone internal router in OSPF, and as in OSPF, the backbone must be contiguous. Level 2 routers communicate via Hellos that are understood only by other Level 2 routers. To fracture the area means the loss of routing information. As in Level 1 routing, the link-state database is identical on all Level 2 routers, although the database contains prefixes of addresses in other areas as opposed to internal area addresses.

The Level 1-2 Router

The router that has everything is, of course, the Level 1-2 router. This is both an intra-area and interarea router. Its characteristics are similar to those of an ABR in OSPF. This router might have neighbors in different areas because it sends both Level 1 and Level 2 Hellos and can thus communicate with everyone. It holds both a Level 1 database for the Level 1 area to which it is connected and a Level 2 database with all the information for interarea routing.

The Level 1-2 configuration is convenient because the router informs Level 1 routers that it is a Level 2 router and can forward traffic to other areas. It can also inform other Level 2 routers of the areas to which it is connected. Although convenient, it consumes more resources on the router in terms of memory and CPU and uses more bandwidth in maintaining the link-state databases. This configuration is the default configuration on Cisco routers.

Basic Principles of Area Routing

Having reviewed the complexities of ISO addressing and the corresponding hierarchical nature of Integrated IS-IS routing, you now need to pull it all together. This section explains the basics of Integrated IS-IS routing. The following list shows what happens when a router receives a packet to forward and how decisions are made in area routing, based on the ISO destination address of the incoming packet:

1. When a router receives traffic to route to a remote destination, it performs the customary routing table lookup.
2. The router strips off the system ID and the SEL to reveal the Area portion of the address. If the Area address is the same as that of the router, it routes the packet toward the host system ID using the Level 1 database.
3. If the Area is different, the router does one of the following:
 - (a) Sends the packets to the nearest Level 2 router (if the router is a Level 1 router).
 - (b) Looks up the route in the forwarding database to find a match, if the router is a Level 2 router.
 - (c) Resolves the address to the longest match, ensuring the greatest level of accuracy.

Summarization, as described in Chapter 2, “IP Addressing,” is used by Integrated IS-IS, and therefore prefix routing can reduce the routing tables and speed up the forwarding process.

Areas in Integrated IS-IS are defined on the link, meaning that the entire router is in the area, requiring the router to define the area and not the interface like in OSPF. Figure 10-3 shows how backbone routers are used to connect different areas.

For the Level 2 routing updates to be exchanged, all the routers capable of sending Level 2 updates must be contiguous; Figure 10-4 shows this.

Figure 10-3 Areas and Backbone Routers

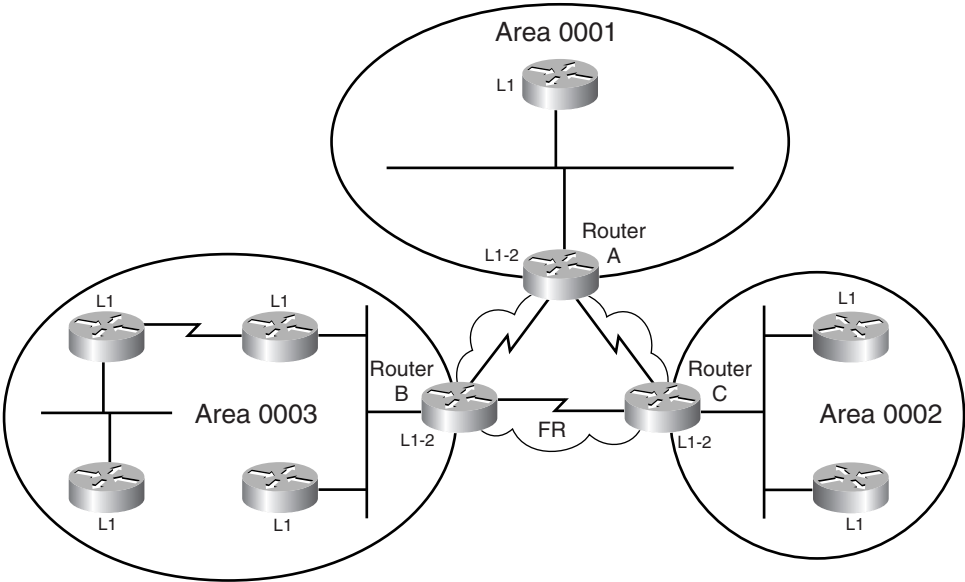


Figure 10-4 A Contiguous IS-IS Backbone

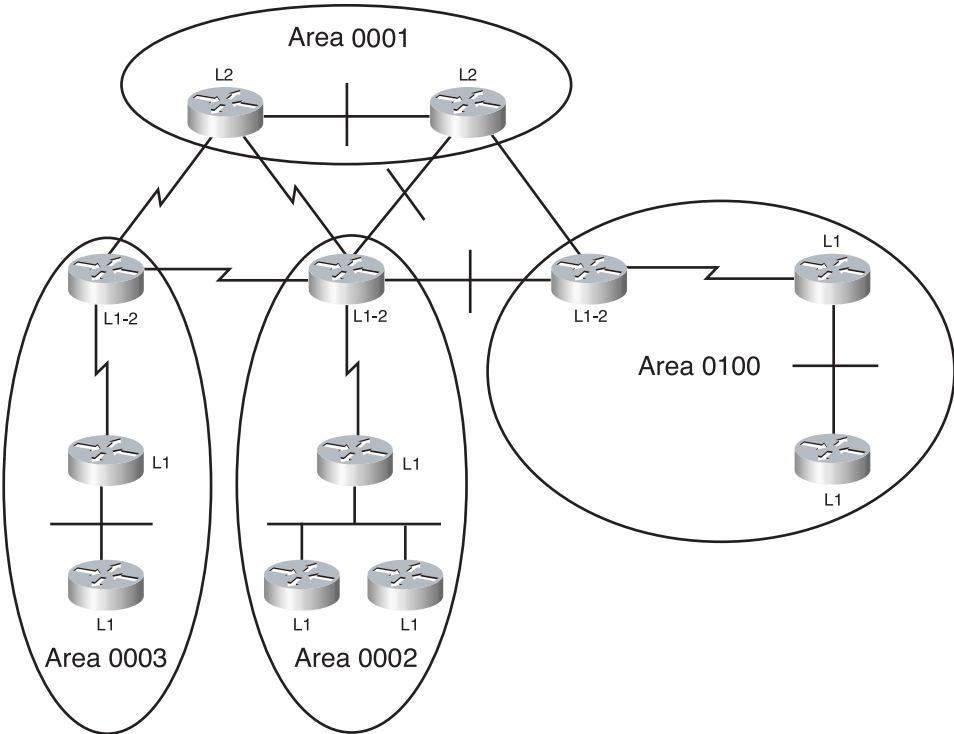
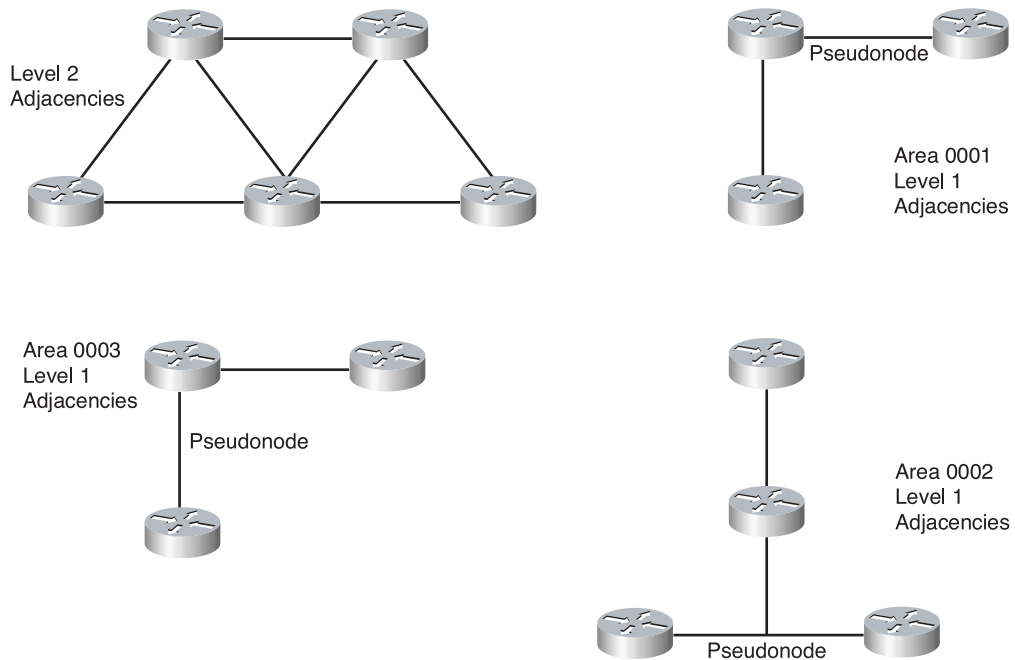


Figure 10-5 *Level 1 and Level 2 Adjacencies for the Topology in Figure 10-4*

Because Level 1 routers communicate only with other Level 1 routers and, likewise, Level 2 routers communicate only between themselves, it can get confusing. Figure 10-5 shows the logical communications or adjacencies set up between the routers shown in Figure 10-4. In Figure 10-5, the area 0001 becomes the backbone area, connecting another area (area 0100). It also shows the communication among L1, L1-2, and L2 routers.

Integrated IS-IS Networks and Interfaces

Routers sharing a common data-link layer become Integrated IS-IS neighbors if the Hello packets that they exchange meet the criteria for forming an adjacency. Although the process of finding a neighbor differs slightly depending on the media, the information sent in the Hellos is essentially the same. Each Hello states the originator of the Hello and the capabilities of its interface. If the Hellos are exchanged and the criteria are met, an adjacency is formed and the Integrated IS-IS neighbors exchange routing information in the form of LSPs. In this way, every router gathers the connected networks of every router to create identical detailed topology maps of the network.

For an adjacency to be formed and maintained, both interfaces must agree on the following:

- The maximum packet size (MTU) that can be transmitted by the interface must be the same.
- Each router needs to be configured as the same level of routing—that is, either Level 1 or Level 2—so that they can decode the Hellos sent by the other router.

- If the routers are both Level 1, they must be in the same area.
- Level 1 routers form adjacencies with each other, and Level 2 routers form adjacencies with other Level 2 routers. For a Level 1 router to communicate with a Level 2 router, one of the routers needs to be configured as a Level 1-2 router. Therefore, to connect to another area at least one of the routers must be configured as a Level 1-2, allowing the Level 2 router to receive the packets destined for another area from the Level 1 router.
- The system ID must be unique to each router.
- If authentication is configured, it must be configured identically on both routers.

The Hello timers (including the holddown timer) must match; otherwise, it could result in a flapping link and endless SPF calculations.

Integrated IS-IS defines two network types—broadcast subnetworks and point-to-point networks—compared with the five types used in OSPF. A broadcast network, as in OSPF, is a multiaccess data link that supports broadcasts and multicasts. The point-to-point links are deemed to be nonbroadcast and can be permanent virtual circuits (PVCs) or dynamically set up switched virtual circuits (SVCs).

The following sections describe adjacencies on point-to-point links, broadcast links, and nonbroadcast multiaccess (NBMA) links.

Establishing Adjacencies on a Point-to-Point Link

A point-to-point link connects two routers. After a Hello packet has been received, each side declares the other side reachable. The routers are neighbors. At this point, each side sends a CSNP. The CSNP is a list of all the links held in the link-state database, which triggers a synchronization of the link-state database on each machine.

Periodic Hellos maintain the adjacency. If a router does not hear a Hello within the Hello holdtime, the router is declared dead and the database is purged of any entries associated with the router. Cisco sets the default Hello multiplier to three. The holdtime is defined as the Hello time multiplied by the Hello multiplier, making the holdtimer expire every 30 seconds.

Establishing Adjacencies on a Broadcast Link

On broadcast links, all the routers running Integrated IS-IS receive packets sent by one router—the DIS—to control the amount of traffic that needs to be generated to maintain the adjacencies and, thus, the databases. The DIS has the responsibility of flooding the LSPs to all connected systems running Integrated IS-IS. More accurately, the DIS floods the LSPs for the pseudonode.

The pseudonode represents the LAN, with each router simulating an interface on the imaginary router. This imaginary router is called the pseudonode. As if it were a real router, the pseudonode

floods a new pseudonode LSP when there is a change in the status of its connections (for example, when a new neighbor comes online).

The adjacencies with the other routers are maintained by the DIS, which sends out Hellos every 3.3 seconds, three times the speed of other routers. This is to ensure the integrity of the adjacencies by identifying a problem very quickly. If there is a problem with the DIS, or a router with a higher priority appears, it is quickly identified and a new router is elected in the place of the old DIS, which is forced into retirement. The election is based on priority and stated on the interface in the event that all the interfaces are set at the default of 64. Then the highest numeric SNPA (data-link address) determines the DIS.

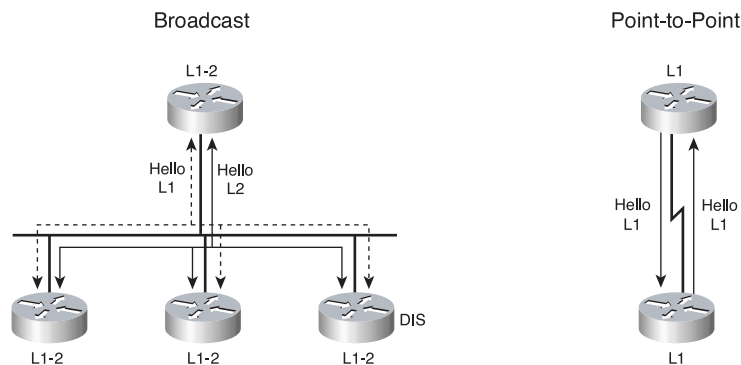
Establishing Adjacencies on an NBMA Link

The creation and maintenance of adjacencies on both point-to-point and multiaccess links is straightforward. It becomes more complicated when the format of the technology cloud allows either link technology to be used.

An NBMA link is neither a broadcast medium nor a point-to-point link; it is a bit of both. Frame Relay, ATM, and X.25 are examples of NBMA. Using PVCs, NBMA provides multiple connections, which could be viewed as a LAN. The confusion occurs when Integrated IS-IS sees the link is multiaccess. Having no knowledge of multiaccess WAN clouds, Integrated IS-IS believes that the medium is some form of LAN and therefore has broadcast capabilities. However, the medium is a WAN, not a LAN, and although the LAN can be simulated, the WAN cloud has no inherent broadcast capabilities.

To avoid complications and possible errors, Cisco recommends that the links be configured as a series of point-to-point links. Figure 10-6 shows a comparison of the broadcast and point-to-point topologies supported by IS-IS.

Figure 10-6 *Network Topologies Supported by IS-IS*



Network Layer Protocols Used in Integrated IS-IS

The PDU is created by the network layer and encapsulated directly into the data-link frame. All the Integrated IS-IS packets share the same eight-octet header. After the fixed header, there are a number of optional variable-length fields that contain specific routing-related information. These variable-length fields are called TLV or CLV.

The fields found in the fixed header of every IS-IS PDU are each one octet in length. Table 10-5 explains the fixed header.

Table 10-5 *The Fixed Header Common to All IS-IS PDUs*

Field	Length of Field in Octets	Description
Intradomain Routing Protocol	1	All Integrated IS-IS PDUs have a value of 0x83. This identifies the packet.
Length Indication	1	States the length of the fixed header.
Version/Protocol ID	1	Set to 1.
ID Length	1	The size of the system ID in the NSAP. This can be an integer between 1 and 8. The Cisco default setting is 6, which is represented by 0, to show that the default has not been changed.
Reserved/Packet Type	1	The first 3 bits are reserved, set to 0, and ignored. The Packet Type indicates whether this is a Hello, LSP, or SNP.
Version	1	Set to 1.
Reserved	1	Set to 0 and ignored.
Maximum Area Addresses	1	States the number of area addresses permitted for this area. Multiple area addresses are configured during transitions. Cisco set the maximum address size to 3, which is represented by 0.

After this common header are the PDU-specific fields and the variable-length fields, creating the different packet types seen in Integrated IS-IS. The PDU-specific fields are discussed in the sections of this chapter relevant to the different packet types: Hellos, LSPs, and SNPs. The variable-length fields are discussed in the section on TLVs.

There are three Integrated IS-IS packets, as the following list describes:

- **Hello** — These packets create and maintain neighbor relationships and adjacencies. There are three types of Integrated IS-IS Hello packet. The type of packet is defined in the fixed header under the Type field and allows the packet to be handed off to the appropriate process. The different types are as follows:
 - **LAN Level 1** — Generated by Level 1 and Level 1-2 routers
 - **LAN Level 2** — Generated by Level 2 and Level 1-2 routers
 - **Point-to-point** — Generated by Level 1, Level 2, and Level 1-2 routers
- **LSP** — LSPs hold information on the neighbors connected to the router. There are two types of LSP, as follows:
 - **Level 1** — Generated by Level 1 and Level 1-2 routers
 - **Level 2** — Generated by Level 2 and Level 1-2 routers
- **Sequence number packet (SNP)** — SNPs describe the LSPs in the transmitting router's link-state database. The information is condensed and is never flooded but only sent between neighbors. SNPs ensure link-state databases synchronization by:
 - Distributing groups of LSPs on a LAN without explicit individual acknowledgements
 - Acknowledging individual LSPs
 - Requesting LSPs at startup

There are two types of SNP for each level of routing, as follows:

- **Complete SNP (CSNP)** — Includes every LSP in the database:
 - Level 1
 - Level 2
- **Partial SNP (PSNP)** — Includes a subset of LSPs, used to request individual LSPs and to acknowledge receipt of these LSPs:
 - Level 1
 - Level 2

The following sections describe Hello packets, LSPs, and SNPs in more detail.

The Format of the Hello Packet

There are three different Hellos, as follows:

- Point-to-point Hello

- LAN Level 1 Hello
- LAN Level 2 Hello

Because the point-to-point and broadcast media work differently, the adjacencies are formed in a different manner. There need to be separate Hellos for the point-to-point networks and the broadcast networks. A point-to-point network has only one other router with which to communicate and, therefore, after the level of routing has been established, the appropriate updates can be sent. However, a broadcast network is a multiaccess network and can have a mixture of both Level 1 and Level 2 routers. For this reason, the broadcast or LAN network has two Hello formats, the Level 1 format and the Level 2 format. The Hellos for the broadcast media are referred to as LAN Hellos.

Point-to-point Hello packets are used over point-to-point links. Table 10-6 shows the point-to-point Hello.

Table 10-6 The Point-to-Point Hello

Field	Length of Field in Octets	Description
Fixed Integrated IS-IS header	8	Common to all Integrated IS-IS PDUs
Circuit Type	1	States whether the transmitting router is Level 1, Level 2, or, if both hellos are received, Level 1-2
Source ID	ID length	The system ID from the NSAP of the transmitting router
Holding Time	2	How long the neighbors must wait for a Hello before they can declare the transmitting router to be dead
Packet Length	2	The length of the Hello packet in octets
Local Circuit ID	1	Identifier for the transmitting interface, which is unique to the transmitting router

LAN Hello packets are used over broadcast links. The PDU is the same for all router types, although the values within the fields differ.

Table 10-7 shows the LAN Hello.

Table 10-7 *LAN Hello Packet Format*

Field	Length of Field in Octets	Description
Fixed Integrated IS-IS header	8	Common to all Integrated IS-IS PDUs.
Circuit Type	1	States whether the transmitting router is a Level 1, Level 2, or, if both Hellos are received, Level 1-2.
Source ID	ID length	The system ID from the NSAP of the transmitting router.
Holding Time	2	How long the neighbors must wait for a Hello before they can declare the transmitting router to be dead.
Packet Length	2	The length of the Hello packet in octets.
Priority	2	Used in the election of the DIS (highest priority wins).
LAN ID	ID length + 1	The DIS uses its system ID plus an additional octet to name the LAN. The extra octet is used to identify the LAN from other LAN connections on the DIS.

The Format of the LSP

The LSP from a Level 1 router is flooded to all routers in the area. The LSP contains a list of all the adjacencies.

Likewise, a Level 2 router floods the LSP to every other Level 2 router in the domain. However, this LSP contains the list of adjacencies to other Level 2 routers and the areas that the transmitting router can reach. The TLVs hold the Level 1 and Level 2 information, allowing the LSP format to be the same for both Level 1 and Level 2 routers.

Table 10-8 shows the format of the LSP.

Table 10-8 LSP Packet Format

Field				Length of Field in Octets	Description
Fixed Integrated IS-IS header				8	Common to all Integrated IS-IS PDUs
Packet Length				2	The length of the entire LSP
Remaining Lifetime				2	How long in seconds before the LSP is purged from the database
LSP ID				ID length + 2	A three-part ID: • 6 octets for the transmitting system ID or pseudonode ID of the transmitting DIS • 1 octet = zero for router 1 octet = nonzero for pseudonode • 1 octet = fragment bit, if the LSP information is too great to be contained in one PDU
Sequence Number				4	Used to determine the latest LSP version
Checksum				2	The checksum on the contents of the LSP
P	ATT	OL	IS Type	1	Together the Partition (P) bit, Attached (ATT) bit, Overload (OL) bit, and IS Type take up 1 octet. P—The Partition bit is used by Level 2 routers to identify whether automatic partition repair is supported. Cisco does not support this feature and the value is always 0. ATT—The Attached bit is used in Level 1 LSPs generated by Level 1-2 routers. It shows Level 1 routers a potential exit from the area. Level 1 routers use this bit to determine the closest Level 2 router. Overload bit (OL)—If the generating router has run out of memory for the link-state database, this bit is set to 1. Other routers treat this router as a host, forwarding packets destined to networks directly connected to routers setting the OL bit. If the router has run out of memory for the link-state database, the database may be incomplete and unreliable as a forwarding router. IS Type is used to indicate whether the router is a Level 1 or Level 2 router.

The Format of the SNP

As with the LSP, the SNP format is the same for both Level 1 and Level 2 routers. The CSNP has two additional fields to the PSNP. The extra fields define the address range, which states all the LSPs that are included in the CSNP. This is to ensure that if the database is so large that several PDUs need to be sent, the information can be reassembled at the receiving router.

Table 10-9 shows the format of the SNP.

Table 10-9 *SNP Packet Format*

Field	Length of Field in Octets	Description
Fixed Integrated IS-IS header	8	Common to all Integrated IS-IS PDUs
Packet Length	2	The length of the entire LSP
Source ID	ID length + 1	The system ID from the NSAP of the transmitting router
Start LSP ID	ID length + 2	Refer to the definition of LSP ID in Table 10-8
End LSP ID	ID length + 2	Refer to the definition of LSP ID in Table 10-8

TLVs

TLVs, sometimes called CLVs, are one of the main strengths of Integrated IS-IS. TLVs provide flexibility and extend the functionality of the protocol. The TLV fields are variable in length and are appended to the various packet formats. This means that the protocol can adapt to the changing needs and advances in technology by defining a new TLV.

The structure of the TLV is as follows:

- **Type or Code** — State the Type of field. This code identifies the TLV and all the characteristics that pertain to it. (For example, TLV 128 defines the capability to carry IP routes in IS-IS packets.) In essence the TLV 128 is Integrated IS-IS.
- **Length** — The length of the following field. This is important because the next field can be any length. By identifying the length of the field, error detection is possible.
- **Value** — The information, whether it is IP routes, IS-IS neighbors, or authentication.

One of the most radical TLVs was the introduction of TLV 128, as defined in RFC 1195 in 1992. This extended Integrated IS-IS to support IS-IS and IP routing. In the future, it is expected that a TLV will be created to support IPv6.

To read the definition of these TLVs and, in particular, the various codes, you must read the standards. ISO/IEC 10589 defines the TLV codes 1 to 10; RFC 1195 defines the TLV codes 128 to 133.

It is important to note which TLVs are supported by your equipment because this determines the design and configuration of the network. The receiving router ignores TLVs that are not supported.

Table 10-10 describes the main TLVs, states the document that defines them, and names the PDUs to which they are appended.

Table 10-10 *TLVs Implemented by Cisco*

TLV	Name	Source Document	Used by PDU	Description
1	Area Addresses	ISO 10589	Hello and LSP Levels 1 and 2	The area addresses configured on the router. Cisco supports a maximum of three area addresses.
2	IS Neighbors	ISO 10589	LSP Level 1 and 2	Lists the neighbors of the transmitting router. The neighbor ID is the system ID plus an extra octet. If the neighbor is a pseudonode, the last octet has a positive value. If the neighbor is a router, the last octet has a value of 0x00. Of the many options within this TLV, the default metric is the only one supported and contains a value between 0 and 63.
3	ES Neighbors	ISO 10589	LSP Level 1	Similar to IS neighbors, assumes the same cost to many ES neighbors as connected via broadcast media. ES neighbors with different costs appear as separate entries.
5	Prefix Neighbors	ISO 10589	LSP Level 2	The same as an ES neighbor but stating an address prefix or domain and area rather than an ES.
6	IS Neighbors	ISO 10589	Hello	Lists all the system IDs of the routers from which a valid Hello has been received. Level 1 routers list Level 1 neighbors; Level 2 routers list Level 2 neighbors.
8	Padding	ISO 10589	Hello	Used to pad a Hello PDU so that it can be a minimum length. The padding is ignored. Cisco uses all 0s to pad packets.
9	LSP Entries	ISO 10589	SNP	The state of an LSP is given via the Remaining Lifetime, LSP ID, Sequence Number, and Checksum. This identifies each LSP and ensures no duplication or corruption.

Table 10-10 *TLVs Implemented by Cisco (Continued)*

TLV	Name	Source Document	Used by PDU	Description
10	Authentication	ISO 10589	Hello, LSP Level 1 and 2, and SNP	The first octet states the type of authentication used. Of the 255 possible values, only one type of authentication—clear text—is defined in ISO 10589. An interface can be configured with a transmit password. If the receiving interface does not receive the expected password, the Hello is rejected.
128	IP Internal Reachability Information	RFC 1195	LSP Level 1 and 2	Lists all IP addresses that the transmitting router knows about from interfaces within the area.
129	Protocols Supported	RFC 1195	Hello	States the protocols the transmitting router supports (whether CLNS only, IP only, or both CLNS and IP).
131	Inter-Domain Routing Protocol Information	RFC 1195	LSP Level 2	Allows information from external routing protocols to be carried in Level 2 LSPs. Similar to route tags in RIPv2, EIGRP, and OSPF.
132	IP Interface Address	RFC 1195	Hello and LSP Level 1 and 2	The IP address or addresses of the transmitting interface.

Foundation Summary

The “Foundation Summary” section of each chapter lists the most important facts from the chapter. Although this section does not list every fact from the chapter that will be on your exam, a well-prepared candidate should, at a minimum, know all the details in each “Foundation Summary” before going to take the exam.

Figure 10-7 shows the format of a CLNS address.

Figure 10-7 An ISO Address

IDP		DSP		
AFI (1 octet)	IDI	High Order DSP	System ID (1-8 octets)	NSEL (1 octet)
AREA			ID	SEL

For routing purposes, ISO 10589 specifies three parts to the address, namely how to get to the area, how to find the host, and how to find the application within the host. However, there are only two elements in the address used to find a destination host: the address that locates the area and the host within that area. Therefore, Integrated IS-IS has two levels of hierarchy, as follows:

- **IDP**—The IDP is used to route to the domain, or autonomous system. The IDP is given by the ISO and identifies the body responsible for assigning the format for the rest of the address by defining the DSP structure. The following two parts comprise the IDP:
 - **AFI (authority and format identifier)** —AFI is the first octet of the address. It is defined as one octet. It states the authority in charge of issuing addresses and the format of the rest of the address, within the constraints of IDI and DSP, to describe Area, ID, and SEL.
 - **IDI (initial domain identifier)** —IDI is the suborganization to the AFI. (As an analogy, the U.S. government is the AFI and the General Services Agency is the IDI, otherwise known as GOSIP.)
- **DSP**—DSP is used to route within the autonomous system. The authority specified by the IDP may further subdivide the DSP and may assign subauthorities responsible for parts of the DSP. The DSP can be broken down into the following fields:
 - **High Order DSP** —High Order DSP is typically the area within the autonomous system.
 - **System ID** —System ID can have a value between one and eight octets. It must have the same value throughout all systems in the autonomous system. Cisco uses six

octets, which is a common solution because it allows the use of the MAC address to autoconfigure the system.

— **NSEL** —NSEL identifies the particular service at the network layer to which to hand the packet.

Characteristics of a Level 1 router include the following:

- An intra-area router
- Similarity to an OSPF stub router
- Knowledge of the network limited to the area
- A link-state database with all the routing information for the area
- The address of the closest Level 2 router to which to send traffic destined for another area
- Neighbors must be in the same area
- A DIS that is elected on LANs

Characteristics of a Level 2 router include the following:

- An interarea router
- Similarity to a backbone internal router in OSPF
- Level 2 routers must be contiguous (that is, the area cannot be fractured)
- Neighbors can be in different areas
- A Level 2 database with all the information for interarea routing

Characteristics of a Level 1-2 router include the following:

- An intra-area router and an interarea router
- Similar to an area boundary router (ABR) in OSPF
- Neighbors can be in different areas
- A Level 2 database with all the information for interarea routing
- A Level 1 database for each area to which it is connected
- Informs Level 1 routers that it is a Level 2 router and can forward traffic to other areas
- Informs other Level 2 routers of the areas to which it is connected

For an adjacency to be formed and maintained, both interfaces must agree on the following:

- The maximum packet size (MTU) that can be transmitted by the interface must be the same.
- Each router needs to be configured as the same level of routing—that is, either Level 1 or Level 2—so that they can decode the Hellos sent by the other router.

- If the routers are both Level 1, they must be in the same area.
- Level 1 routers form adjacencies with each other, and Level 2 routers form adjacencies with other Level 2 routers. For a Level 1 router to communicate with a Level 2 router, one of the routers needs to be configured as a Level 1-2 router. Therefore, to connect to another area, at least one of the routers must be configured as a Level 1-2, allowing the Level 2 router to receive the packets destined for another area from the Level 1 router.
- The system ID must be unique to each router.
- If authentication is configured, it must be configured identically on both routers.

There are three Integrated IS-IS packets, as the following list describes:

- **Hello** — These packets create and maintain neighbor relationships and adjacencies. There are three types of Integrated IS-IS Hello packet. The type of packet is defined in the fixed header under the Type field and allows the packet to be handed off to the appropriate process. The different types are as follows:
 - **LAN Level 1** — Generated by Level 1 and Level 1-2 routers
 - **LAN Level 2** — Generated by Level 2 and Level 1-2 routers
 - **Point-to-point** — Generated by Level 1, Level 2, and Level 1-2 routers
- **LSP** — LSPs hold information on the neighbors connected to the router. There are two types of LSP, as follows:
 - **Level 1** — Generated by Level 1 and Level 1-2 routers
 - **Level 2** — Generated by Level 2 and Level 1-2 routers
- **Sequence number packet (SNP)** — SNPs describe the LSPs in the transmitting router's link-state database. The information is condensed and is never flooded but only sent between neighbors. SNPs ensure link-state databases synchronization by:
 - Distributing groups of LSPs on a LAN without explicit individual acknowledgements
 - Acknowledging individual LSPs
 - Requesting LSPs at startup

There are two types of SNP for each level of routing, as follows:

- **Complete SNP (CSNP)** — Includes every LSP in the database:
 - Level 1
 - Level 2

— **Partial SNP (PSNP)** —Includes a subset of LSPs, used to request individual LSPs and to acknowledge receipt of these LSPs:

Level 1

Level 2

There are three different Hellos, as follows:

- Point-to-point Hello
- LAN Level 1 Hello
- LAN Level 2 Hello

Q&A

The following questions test your understanding of the topics covered in this chapter. After you have answered the questions, check your answers in Appendix A. If you get an answer wrong, review the answer and ensure that you understand the reason for your mistake. If you are confused by the answer, refer to the appropriate text in the chapter to review the concepts.

1. Which system generates the pseudonode?
2. What is a CSNP? When is it used?
3. What is a PSNP? When is it used?
4. A new router comes online on a multiaccess link, and the priority is the same as that of the DIS for the segment. What action is taken?
5. What happens in the event of the DIS dying?
6. Integrated IS-IS can be used to send information about which routed protocols?
7. How often does the DIS send out a Hello packet?
8. What is the name of the link-state algorithm used to create entries for the routing table?
9. What is the relationship between the Hello timer and when the path is considered to have died?
10. Integrated IS-IS areas are similar to which type of areas in OSPF?
11. Describe one design restriction in configuring Level 2 routing.
12. Given the following address:
49.0001.2222.2222.2222.00
Is this a NET or NSAP address? Give reasons for your choice.
13. What is a pseudonode and what is its purpose?
14. State two reasons why a router might not be able to find a neighbor.
15. Explain briefly why two routers cannot have the same system ID within the area.
16. What does TLV stand for? Briefly explain its purpose.
17. How many link-state databases does a Level 1-2 router have?
18. Integrated IS-IS packets run directly on top of which layer?
19. What is the NET address associated with in the IS-IS addressing scheme?
20. Describe briefly the main characteristics of a Level 1 IS.



This chapter covers the following topics, which you need to understand to pass the CCNP/CCDP/CCIP BSCI exam:

- Integrated IS-IS operation
- Integrated IS-IS design considerations

Integrated IS-IS Protocol Operation

The topics in this chapter describe the routing protocol Integrated IS-IS by explaining its operation and design considerations. The detailed explanation of the routing protocol assumes knowledge of the general operation and purpose of routing protocols and, in particular, link-state routing protocols.

Because Integrated IS-IS is similar to the Open Shortest Path First (OSPF) protocol, you should read the chapters on OSPF and reinforce the fundamentals of link-state protocols.

“Do I Know This Already?” Quiz

The purpose of the “Do I Know This Already?” quiz is to help you decide what parts of this chapter to use. If you already intend to read the entire chapter, you do not necessarily need to answer these questions now.

The 12-question quiz, derived from the major sections in the “Foundation Topics” portion of the chapter, helps you to determine how to spend your limited study time.

Table 11-1 outlines the major topics discussed in this chapter and the “Do I Know This Already?” quiz questions that correspond to those topics.

Table 11-1 “Do I Know This Already?” Foundation Topics Section-to-Question Mapping

Foundation Topics Section	Questions Covered in This Section
Integrated IS-IS Operation	1–10
Integrated IS-IS Design Considerations	11–12

NOTE The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark this question wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

1. What triggers an LSP update to be flooded to neighbors?
 - a. When an adjacency goes up or down
 - b. When there is a change in the state of the interface
 - c. When there is a change in the routing table
 - d. When an LSP is received
2. If an LSP that already exists in the database is received, the received LSP is ignored on what condition?
 - a. It is a valid packet that is older than the one in the database.
 - b. It is a valid packet with the same sequence number as the one in the database.
 - c. It is not a valid packet because it is corrupted.
 - d. It is a valid packet that is newer than the one in the database.
3. A point-to-point link uses which of the following packets to create and maintain neighbor tables in addition to the link-state databases?
 - a. PSNP
 - b. Hellos
 - c. CSNP
 - d. LSP
4. What is the PATH table used for?
 - a. Each candidate route is placed in the PATH database while the metric is examined to determine if it is the shortest path to the destination.
 - b. The forwarding database examines the PATH table to ascertain the details of the chosen path.
 - c. Private Allocation Tree Hidden (PATH table) lists the RFC 1918 addresses to prevent them leaking into the Internet.
 - d. It holds the LSP immediately prior to flooding.

5. There are many criteria used in selecting routes. These criteria are used in a certain order to ensure that the best decision has been made. Which of the following criteria is used first in the selection process?
 - a. Longest match, that is, the route with the most specific subnet mask
 - b. Level 1 routes
 - c. Internal routes (within the autonomous system)
 - d. Routes with ToS defined
6. What is the default cost applied to the outgoing interface of a Cisco router configured to run IS-IS?
 - a. 15
 - b. 10
 - c. Inverse of bandwidth
 - d. There is no default. It must be configured.
7. At what point is the forwarding database created?
 - a. After all the LSPs are received
 - b. As soon as the LSP is received
 - c. After the SPT is built
 - d. Once the neighbor table is built
8. When the router receives user data and error reports, to which process are they sent?
 - a. The decision process
 - b. The forwarding process
 - c. The receive process
 - d. The update process
9. When the router receives hellos or routing information (in the form of SNPs and LSPs) to which process are they sent?
 - a. The decision process
 - b. The forwarding process
 - c. The receive process
 - d. The update process

10. Which of the following are IS-IS rules for summarization?
- a. All Level 1 routers can summarize routes.
 - b. All Level 2 routers can summarize at the area boundary.
 - c. If a Level 1-2 router is summarizing routes sent to a Level 2 router, all Level 1-2 routers must summarize in the same way.
 - d. All the above.
11. What is the advantage of designing a totally flat network running Level 1-2 routing on every router?
- a. This design allows for easy migration to multiple areas.
 - b. This design allows for easy system administration.
 - c. This design requires fewer resources of either the network or the router.
 - d. Summarization can be configured within the area.
12. What could result if the Level 1-2 router fails?
- a. Another router would be elected as the Level 1-2 router.
 - b. The autonomous system would not be able to communicate with the outside world.
 - c. The area would be completely annexed from the rest of the network.
 - d. No summary routes would be sent into the backbone area.

The answers to this quiz are found in Appendix A, “Answers to Chapter ‘Do I Know This Already?’ Quizzes and Q&A Sections.” The suggested choices for your next step are as follows:

- **6 or less overall score** — Read the entire chapter. This includes the “Foundation Topics” and “Foundation Summary” sections, and the “Q&A” section at the end of the chapter.
- **7–9 overall score** — Begin with the “Foundation Summary” section, and then go to the “Q&A” section at the end of the chapter. If you have trouble with these questions, read the appropriate sections in “Foundation Topics.”
- **10 or more overall score** — If you want more review on these topics, skip to the “Foundation Summary” section, and then go to the “Q&A” section at the end of the chapter. Otherwise, move to the next chapter.

Foundation Topics

Integrated IS-IS Operation

Before you can fully grasp the details of IS-IS and how it routes traffic, you must understand the routing protocol operation. This operation is similar to the section on IS-IS convergence detailed in Chapter 5, “IP Link-State Routing Principles.” The section in Chapter 5 dealt with how the routing protocol identifies a change in the network and updates the routing tables. This section describes how all the databases for IS-IS are created and maintained, not simply how they are updated.

A high-level overview of the Integrated IS-IS operation follows:

1. Routers send Hellos out of all Integrated IS-IS interfaces to discover neighbors and to form adjacencies.
2. Routers sharing a common data link become neighbors.
3. Routers build link-state packets (LSPs) based on local Integrated IS-IS interfaces and prefixes learned from other adjacent routers. These are sent to all neighbors.
4. Routers flood received LSPs to all adjacent routers except to the neighbor from which the LSP was received.
5. When new or different LSPs are received, the router adds the LSPs to the link-state database.
6. The router calculates the SPF for each destination and constructs the shortest path tree (SPT) and the forwarding database.

The routing process for IS-IS is divided into four stages:

- Update
- Decision
- Forwarding
- Receive

The following sections describe each stage in detail.

The Update Process

The router can forward data packets to the remote destination only if it has an understanding of the network topology. The router can then determine the best path to the remote network and send the data packet on the next stage of its journey.

To ensure that every router has the information required, each router generates its own LSP. The LSP lists the router's neighbors, which are propagated throughout the network. The flooding of LSPs ensures every router has the same set of LSPs and that the link-state databases are identical.

LSPs are generated whenever there is a change in the network, often because of a configuration change. However, any of the following instances trigger a new LSP to be flooded throughout the network:

- An adjacency either comes up or down (for example, a new router comes online).
- An interface on the router changes state or is assigned a new metric.
- An IP route changes (for example, because of redistribution).

During the flooding process, a router both receives and sends LSPs. The action of each instance has clear rules of engagement.

Sending and Receiving an LSP

On receipt of the new LSP, the router stores it in the link-state database and marks it for flooding. If the LSP is already present in the database, the router just acknowledges and ignores it. The router sends the duplicated LSP to its neighbors, which in turn flood to their neighbors and onward to the very edges of the network. Because Level 1 and Level 2 routers have their own link-state databases, Level 1 LSPs are flooded throughout the area; Level 2 LSPs are sent across all Level 2 adjacencies.

The process of propagating LSPs differs slightly, depending on which medium the LSP was received. The following sections describe the propagation process for point-to-point and broadcast links.

Propagating LSPs on a Point-to-Point Interface

The point-to-point link does not need to ensure that multiple systems have synchronized databases. With only one other router with which to work, some reliance is given to the router's capability to determine the need to update so that bandwidth can be optimized.

The following list describes the flooding process on a point-to-point link:

1. When an adjacency is established, both sides send a complete sequence number packet (CSNP) with a compressed version of their link-state database (Router ID and the sequence number).
2. If there are any LSPs in the receiving router's database that were not in the CSNP it received, it sends a copy of the missing LSPs to the other router.
3. Likewise, if the database is missing any LSPs received in the CSNP, the receiving router requests the detailed or full LSP to be sent.

4. The individual LSPs are requested, sent, and acknowledged via partial-sequence number packet (PSNPs).
5. When an LSP is sent, the router sets a timer, and if no explicit acknowledgement has been received before the timer expires, the LSP is resent. This timer is the **minimumLSPTransmission-interval** and can be configured; the default on a Cisco router is five seconds.

Propagating LSPs on a Broadcast Link

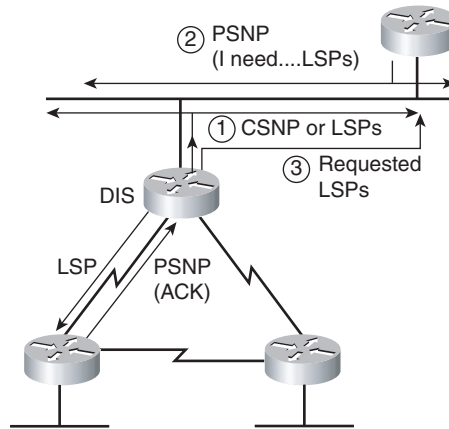
A broadcast link may need to send out both Level 1 and Level 2 updates. The link sends these updates using multicast MAC addresses to all Level 1 routers and all Level 2 routers. Because the pseudonode is just that, a pretend system, a real node or system must enter the charade and perform the tasks of the pseudonode. The designated intermediate system (DIS) takes on much of the responsibility for synchronizing the databases on behalf of the pseudonode. The DIS has three tasks, as follows:

- Creating and maintaining adjacencies
- Creating and updating the pseudonode LSP
- Flooding the LSPs over the LAN

The following lists the main steps in the flooding process:

1. On receipt of a CSNP, the router compares each compressed LSP with the link-state database.
2. If the database has a newer version of the LSP sent in the CSNP, or if there is no instance of the LSP in the CSNP, the router multicasts the LSP onto the LAN.
3. If the database is missing an LSP that was in the CSNP, it sends a PSNP requesting the full LSP. Although the router multicasts, it is only the DIS that takes action.

Figure 11-1 summarizes the flow of CSNPs and PSNPs on broadcast and point-to-point links.

Figure 11-1 *The Propagation of CSNPs and PSNPs*

Determining Whether the LSP in the Database Is Valid

The LSP contains three fields that help determine whether the LSP that has been received is more recent than that held in the database, and whether it is intact or has been corrupted. These three fields are as follows:

- **Remaining Lifetime** — This is used to age-out old LSPs. If an LSP has been in the database for 20 minutes, it is assumed that the originating router has died. The refresh timer is set to 15 minutes.
If the lifetime expires, the LSP has the content removed, leaving only the header. The lifetime is set to show that it is a new LSP, and then it is flooded through the network. All receiving routers accept the mutilated LSP, recognize that this means the route is bad, and purge the existing LSP from their databases.
- **Sequence Number** — This is an unsigned 32-bit linear number. The first LSP is allocated the sequence number 1, and the following LSPs are incremented by 1.
- **Checksum** — If a router receives an LSP and the checksum does not compute correctly, the LSP is flushed and the lifetime is set to 0. The router floods the LSP, all routers purge the LSP, and the originating router retransmits a new LSP.

The Decision Process

After the link-state databases have been synchronized, it is necessary to decide which path to take to reach the destination. Because the routers and hosts may have multiple connections to each other, there may be many paths from which to choose.

To make the decision as to the best path, link-state protocols employ the algorithm defined by Dijkstra. This algorithm is extremely complex, but its core function is to create a tree that mirrors the shortest paths to all destinations. Each router builds a shortest path tree with itself as the root. This is achieved by taking all the LSPs from the link-state database and using the Dijkstra algorithm to create the shortest path tree (SPT). The SPT is used in turn to create the forwarding table, also known as the routing table.

Several tables are created during this procedure. The PATH table is the SPT during construction, and the TENT is a tentative database (a scratchpad).

The steps by which the router's SPT and the forwarding database are built are as follows:

- Step 1** The router places itself at the root in the PATH table.
- Step 2** The SPF takes each LSP from the link-state database and selects the best or shortest path by considering the metric for each path in turn. The lowest metric is selected as the shortest path.
- Step 3** The decision process looks at LSP for the node just placed in the PATH table. The metric cost of getting to the node is added to the cost shown in LSP (the cost from the node to the destination advertised). If this path is not already in PATH or TENT with a better cost, put it in TENT.
- Step 4** If the TENT table is empty, stop.
- Step 5** If the TENT table is not empty, find the lowest-cost entry and move the PATH to the TENT table.
- Step 6** The router then returns to step 2 and repeats the remaining steps.

If there is more than one path to a remote destination, the criteria by which the lowest cost paths are selected and placed in the forwarding database are as follows:

1. If there is more than one path with the lowest value metric, Cisco equipment places up to six equal-cost paths into the table. The default number of equal-cost paths is four.
2. Optional metrics are chosen before the default metric, but because Cisco supports only the default metric, this is a moot point.
3. Internal paths are chosen before external paths, because going outside the autonomous system is likely to be a suboptimal route and might be the result of a routing loop.
4. Level 1 paths within the area are more attractive. If the path is within the area, not only is it more efficient to route directly to it, but also going outside the area and returning can be the cause of a routing loop, demanding greater resources and time.

5. The address with the longest match or most specific address in IP is the address with the longest IP subnet mask. This ensures that the closest router is chosen, because prefix routing is configured by summarization that can occur only on area boundaries.
6. Type of Service (ToS) is a field in the IP header that allows for complex routing decisions to be made by adding another level of criteria to the decision process. If ToS has been configured, the path with ToS is chosen over those paths with no ToS.
7. If multiple paths have ToS configured, the path with the shortest route is selected.
8. If the ToS is the same, up to six entries are placed into the forwarding database and the routers load balance between them.
9. If there is no path, the forwarding database sends the packet to the nearest Level 2 router, which is the default router.

The metric defines the cost of the path. Integrated IS-IS has four metrics, only one of which is required and supported. The metrics defined in ISO 10589 are as follows:

- **Default** — Sometimes referred to as cost. Every Integrated IS-IS router must support this metric. Cisco set the default for all interfaces to be 10.
- **Delay** — This optional metric reflects the transit delay.
- **Expense** — This optional metric reflects the monetary expense of the network.
- **Error** — The reliability of the path is determined as the metric.

The ISO metrics are configured on the outgoing interface using an integer value between 0 and 63. A 10-bit field describes the total path to a destination, allowing a value between 0 and 1023. Considering this to be inadequate, Cisco increased the metric size to be 24 bits. The default configuration is using the 6-bit or narrow metric. This metric is the original metric that acquired the name *narrow* only after the wide metric was defined.

In determining the shortest path, the lowest metric is chosen, internal paths are chosen over external paths, and Level 1 routes have precedence over Level 2 routes.

The default or cost metric is the only metric supported by Cisco, because each metric used in Integrated IS-IS requires a different link-state database calculation for both the Level 1 and Level 2 routes.

Other rules have little to do with the decision process but are responsible for ensuring the LSPs are valid. Essentially, Integrated IS-IS ignores incomplete fragments. Any fragmented packet entering the router is ignored if there is no starting fragment with the packet header (because this indicates packet loss or corruption). The O fragment also defines whether the OL bit is set, the Integrated IS-IS type, and whether there are entries in the Area Address options field. Although this does not eliminate the reassembly of corrupted packets, it certainly reduces the load.

The Forwarding and Receive Processes

The IP subnets are carried in the LSPs in the IP Reachability TLV (code 128) and are treated in the same manner as the OSI addresses. The decision process for building the SPT is the routing decision.

After the SPT has been built, the forwarding database can be created. The forwarding table is essentially a lookup table for the longest match, where load balancing occurs on multiple equal-cost paths. The forwarding table for Integrated IS-IS is more relevant to CLNS than to IP because the IP routing information is entered directly into the IP routing table, where IP routes are leaves on the tree of IS-IS.

Although the forwarding database is concerned solely with forwarding the data, it selects the longest match and load balances across equal-cost links.

If the frame is valid, the receive process passes user data and error reports to the forwarding process, whereas routing information Hellos, LSPs, and SNPs are sent to the update process.

The receive process is not described in depth here because it is primarily concerned with CLNS routing and does not concern IP. A detailed description is given in the ISO 10589 standards document.

Integrated IS-IS Design Considerations

Optimizing the network depends on careful planning and design. Although each network is constrained by physical and technical limitations, you must strive to design your network to meet the needs of its users and accommodate the demands of various applications.

In Integrated IS-IS, the fundamental design considerations are areas and addressing. The next sections describe the Integrated IS-IS design considerations of area design, route summarization, and nonbroadcast multiaccess (NBMA) modeling solutions.

Area Design of Integrated IS-IS Routers

When designing a network, you have to compromise. Typically, the trade-off is between reliability and speed. What is most efficient for the network is determined by the requirements of the network and the resources available.

In designing Integrated IS-IS networks and the hierarchical design, you need to consider the data flow in addition to the resources required by the routing protocol. Tuning the update process might be sufficient, although this results in the compromise of resources and reliability. If you reduce the update timers, the databases converge more quickly, but the network could be depleted of necessary resources to route data.

In area design, two SPF databases need to be maintained, which requires the use of additional resources for those routers that straddle more than one area, that is, Level 1-2 routers. The increased resources required are not only in terms of CPU and memory, but also in bandwidth for the propagation of LSPs.

Some typical designs include the following:

- A totally flat network that uses only Level 1 routing. This design will not scale because any change in the network requires a flood of LSPs to all routers, which consequently run the SPF algorithm. However, this simplified design has some advantages: There will be only one SPF database and no problem of suboptimal routing between areas.
- A totally flat network using only Level 2 routing. As the network expands, Level 1 areas may be added. The Level 2 area has complete internal knowledge with the streamlined advantages of one SPF instance. The Level 1 area may well be a company connecting to the ISP, in which case the areas would be added as new customers came online.
- A totally flat network using the Cisco default of every router running Level 1-2 routing. This allows for easy migration to a hierarchical design and overcomes the problem of suboptimal routing. However, this design requires resources to maintain two SPF databases.
- A hierarchical network where the core is running Level 2 routing, with Level 1 areas connecting into the core. Level 1-2 routers are used to connect the areas. Although this is an excellent design, using the strengths of IS-IS, there are still concerns that should be considered. This design results in the use of additional resources and the possibility of suboptimal routing. Configuring the metric for the outbound interface determines the routing decision. This requires a careful knowledge of the topology of the network to ensure that the problem is not compounded. Recent developments have allowed for route leaking to be created, which allows the Level 2 routers to pass some specific routing information into Level 1. This facilitates the optimal routing decision to be selected.

By default, Cisco routers running Integrated IS-IS operate as Level 1-2 routers. You can configure the router to reflect a true Level 1 and Level 2 design, but the default operation overcomes many potential pitfalls, including the following:

- **Area partitions** — In hierarchical routing between areas, if there is a loss of connection between the Level 1 router and the Level 2 router, the subarea could be completely annexed unless there is an alternative route.
- **Loss of data** — If the area becomes partitioned, a section of the network becomes unavailable through a loss of connectivity.
- **Suboptimal routing decisions** — Level 1 areas have knowledge only of networks within their own areas. To reach another area, packets are sent to the nearest Level 2 router. Without additional configuration, the Level 1 router determines the nearest Level 2 router to be the one

with the lowest metric, which translates to the lowest hop count. The metrics used are the default metric of 10, regardless of the bandwidth.

Route Summarization

Summarization has many advantages. It reduces the resource needs of the network and hides network problems within an area. If the router is unaware of a change or problem in the network, the databases are not updated or recalculated, reducing the resources required for SPF calculations. The more details the router knows about the network, the more the router must do to maintain an accurate knowledge of that network; that is, anytime it fails to hear an LSP, however remote and small its detail of the network, the entire network must flood LSPs and recalculate the SPT. Summarization allows the areas to manage the internal knowledge of the network and to summarize that knowledge across area boundaries. The summarization command is given in Chapter 12, “Configuring Integrated IS-IS.”

The rules for summarization using OSPF in multiple areas apply to Integrated IS-IS. The rules for summarizing IP routes in Integrated IS-IS are as follows:

- Level 1-2 routers can summarize the routes held within their area. The summarized routes are propagated to a Level 2 router. This is an efficient method of establishing prefix routing into other areas. The summarization is configured at the Level 1-2 router at the edge of the area. This is similar to the OSPF summarization by the ABR.
- If one Level 1-2 router has route summarization configured, you must configure route summarization on every Level 1-2 router injecting updates into Level 2. If one router is advertising without summarization, all traffic destined for the area is sent to that router because Integrated IS-IS uses longest-match routing. Therefore, the router advertising the more specific route is the chosen one. This is true of Level 1-2 routers sharing the same Level 1 area.
- Level 1 routes cannot be summarized within the area because this is not permitted by the protocol.

Integrated IS-IS NBMA Modeling Solutions in a Switched WAN Network

In the good old days, eight or more years ago, most organizations had a broadcast network (such as Ethernet) for the LAN and point-to-point links (such as a serial link) for the WAN. The broadcast link was multiaccess and the point-to-point link connected two systems, typically over a very low bandwidth link. Accommodating these different technologies when building a routing protocol made sense, but technologies have developed in complexity since then, particularly in the WAN environment.

Frame Relay and ATM are examples of NBMA networks. NBMA networks are not accommodated in Integrated IS-IS. OSPF has a point-to-multipoint configuration option, but Integrated IS-IS does not. The options are to use a multipoint configuration (which equates to a full mesh using broadcast technology), point-to-point subinterfaces, or a combination of both. The recommended solution is point-to-point subinterfaces, although it is possible to configure the interfaces as multipoint.

Foundation Summary

The “Foundation Summary” section of each chapter lists the most important facts from the chapter. Although this section does not list every fact from the chapter that will be on your exam, a well-prepared candidate should, at a minimum, know all the details in each “Foundation Summary” before going to take the exam.

Figure 11-2 shows how IS-IS selects the best path to be entered into the forwarding table.

The routing process is divided into four stages:

- The update process
- The decision process
- The forwarding process
- The receive process

Some typical area designs include the following:

- A totally flat network that uses only Level 1 routing.
- A totally flat network using only Level 2 routing. As the network expands, Level 1 areas may be added.
- A totally flat network using the Cisco default configuration of every router running Level 1-2 routing.
- A hierarchical network where the core is running Level 2 routing with Level 1 areas connecting into the core. Level 1-2 routers are used to connect the areas.

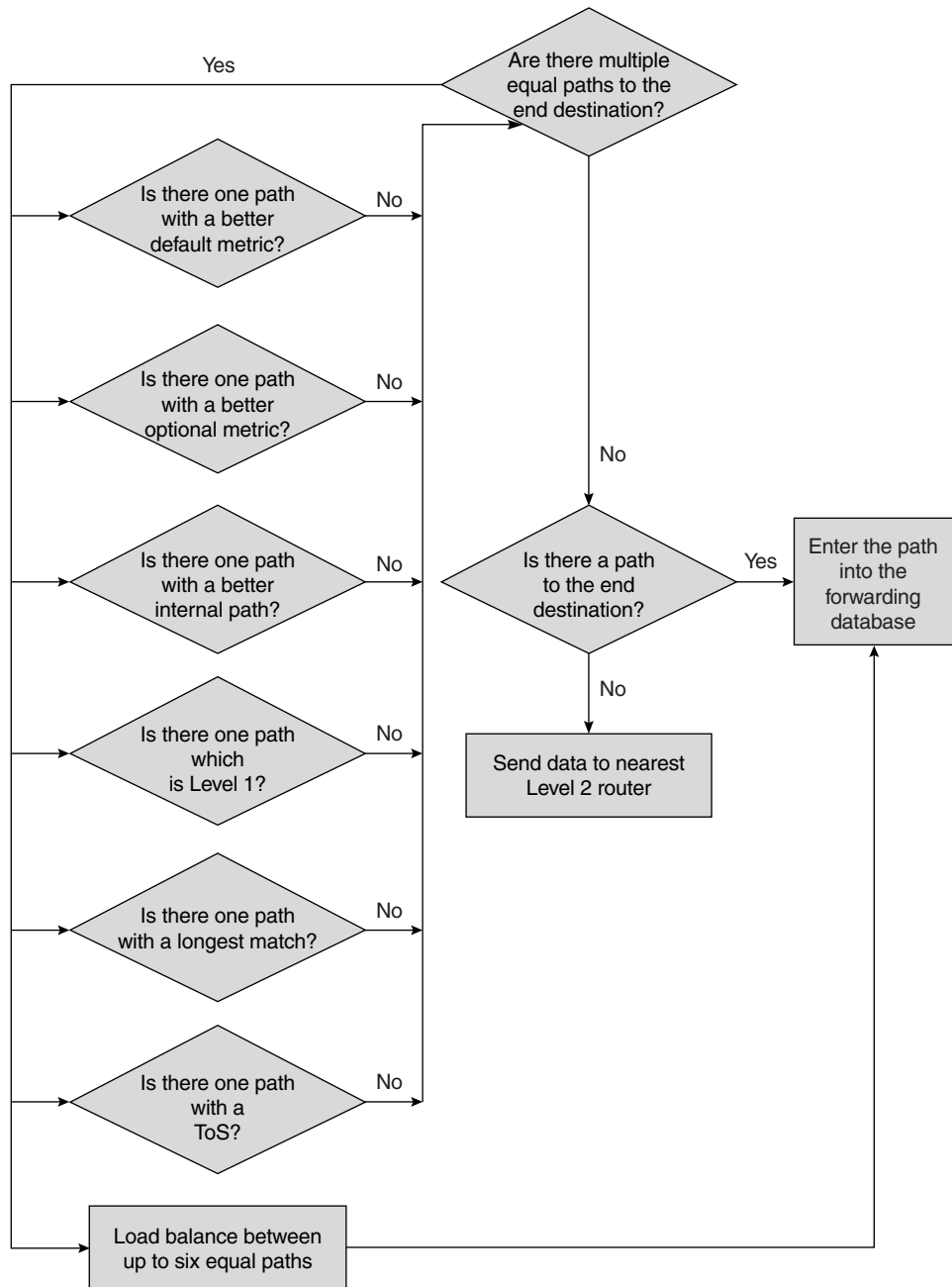
By default, Cisco routers running Integrated IS-IS operate as Level 1-2 routers, which overcomes many potential pitfalls, including the following:

- Area partitions
- Loss of data
- Suboptimal routing decisions

However, this configuration consumes far more resources than a topology that uses a hierarchical design with Level 1, Level 2, and Level 1-2 routers.

Integrated IS-IS does not have a point-to-multipoint configuration option. The alternatives are as follows:

- Configure the link as multipoint, creating a broadcast environment that requires the election of a DIS.
- Configure the interfaces with subinterfaces that are point-to-point.

Figure 11-2 *Selecting the Best Path to Be Entered into the Forwarding Table*

Q&A

As mentioned in the introduction, “All About the CCNP, CCDP, and CCIP Certifications,” you have two choices for review questions. The questions that follow next give you a bigger challenge than the exam itself by using an open-ended question format. By reviewing now with this more difficult question format, you can exercise your memory better and prove your conceptual and factual knowledge of this chapter. The answers to these questions are found in Appendix A.

For more practice with examlike question formats, including questions using a router simulator and multichoice questions, use the exam engine on the CD-ROM.

1. Which systems would you configure as Level 1-2 systems?
2. Which IS-IS configuration uses a full mesh and simulates a broadcast technology?
3. What are the four stages of the routing process?
4. What does an LSP contain?
5. When are LSPs generated?
6. State at least one of the main steps of the flooding process on a point-to-point link.
7. Which three fields determine whether the LSP is valid?
8. Once the link-state databases are synchronized, the Dijkstra algorithm is run. Describe where the router places itself in the tree.
9. State two criteria in determining which paths are to be placed in the forwarding database.
10. What are the ISO metrics?
11. How many equal-cost paths is it possible to have in the IS-IS routing table of a Cisco router?
12. What is a narrow metric?
13. Where is the IS-IS metric applied?
14. What action will the routing process take if it sees an incomplete LSP fragment?
15. Why is the IS-IS default of cost the only metric supported by Cisco?
16. When designing a network for fast convergence, what is the compromise that you need to consider?
17. What is a suboptimal routing decision?
18. Where does route summarization take place?
19. When is a DIS elected in a WAN environment?
20. Explain briefly how the IS-IS NBMA cloud is different than the configuration of the OSPF cloud.



This chapter covers the following topics, which you need to understand to pass the CCNP/CCDP/CCIP BSCI exam:

- Basic configuration of Integrated IS-IS
- Optional Integrated IS-IS commands
- Verifying the Integrated IS-IS operation
- Troubleshooting the Integrated IS-IS operation

Configuring Integrated IS-IS

As in all configurations, the essential commands give a basic configuration, and the more advanced commands either tune the system for efficiency or provide additional configurations for different situations. There is the cake, and then there is the icing. The first section of this chapter deals with the required commands. The next section covers the optional commands. You then learn how to confirm the configuration and how to troubleshoot that configuration to maintain the smooth operation of the network.

The topics in this chapter detail the steps to configuring the IS-IS protocol for integrated routing on a Cisco network. This chapter assumes knowledge of routing protocols—in particular, link-state routing protocols—and the terminology, concepts, and operation of IS-IS.

“Do I Know This Already?” Quiz

The purpose of the “Do I Know This Already?” quiz is to help you to decide what parts of this chapter to use. If you already intend to read the entire chapter, you do not necessarily need to answer these questions now.

The 12-question quiz, derived from the major sections in the “Foundation Topics” portion of the chapter, helps you to determine how to spend your limited study time.

Table 12-1 outlines the major topics discussed in this chapter and the “Do I Know This Already?” quiz questions that correspond to those topics.

Table 12-1 “Do I Know This Already?” Foundation Topics Section-to-Question Mapping

Foundation Topics Section	Questions Covered in This Section
Basic Configuration of Integrated IS-IS	1–3
Optional Integrated IS-IS Commands	4–6
Verifying the Integrated IS-IS Operation	7–10
Troubleshooting the Integrated IS-IS Operation	11–12

NOTE The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark this question wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

1. Which of the following is the correct command to start the Integrated IS-IS routing process?
 - a. **ip router isis**
 - b. **router isis**
 - c. **routing ip isis**
 - d. **router clns**
2. Where is the Integrated IS-IS routing process started?
 - a. At the executive level
 - b. At the interface level
 - c. At both the executive and interface levels
 - d. Underneath the IP routing process
3. What is the purpose of the **net** command?
 - a. To define the summarized address range on the router interface
 - b. To define the area into which the interface is to be placed
 - c. To define the IS-IS address on the interface
 - d. To define the CLNS address for the router
4. Where is the routing level changed from the default of Level 1-2?
 - a. At the executive level
 - b. At the interface level
 - c. Underneath the routing process or at the interface level
 - d. Underneath the IP routing process
5. Which of the following are valid commands for changing the routing level?
 - a. **isis level-2**
 - b. **isis circuit-type level-1**

- c. **isis router level-1**
 - d. **ip router level 1**
6. Which of the following commands summarizes the subnets 140.100.104.0 to 140.100.111.0 at the area boundary?
- a. **summary-address 140.100.104.0 255.255.248.0.**
 - b. **summary-address 140.100.104.0 0.0.7.255.**
 - c. **ip isis summary address 140.100.104.0/21.**
 - d. None of the above; summarization is supported only in OSPF and EIGRP.
7. Which of the following are displayed in the command **show clns neighbor** ?
- a. The contents of the neighbor table
 - b. The routing level as defined at the interface level
 - c. The System ID of the transmitting neighbor
 - d. All of the above
8. Which of the following are displayed in the command **show clns interface** ?
- a. The number of LSPs received on the interface
 - b. The parity check on the received hellos
 - c. The metric of the outgoing interface
 - d. The round trip delay
9. Which of the following are displayed in the command **show isis database** ?
- a. The root of the SPF tree
 - b. The LSPs in the local database
 - c. Whether an LSP has been fragmented
 - d. The sequence number of the LSPs
10. Which of the following commands shows whether the ATT bit has been set?
- a. **show isis hello packets**
 - b. **show isis database**
 - c. **debug isis interface**
 - d. **debug clns interface**

11. Which of the following are displayed in the command **debug isis update-packets** ?
- a. Hellos
 - b. LSPs
 - c. CSNPs
 - d. PSNPs
12. Which of the following are displayed in the command **debug isis adjacency-packets** ?
- a. The population of the PATH database
 - b. The LSPs
 - c. Changes in the state of the adjacencies
 - d. Hello packets

The answers to this quiz are found in Appendix A, “Answers to Chapter ‘Do I Know This Already?’ Quizzes and Q&A Sections.” The suggested choices for your next step are as follows:

- **6 or less overall score** — Read the entire chapter. This includes the “Foundation Topics” and “Foundation Summary” sections, the “Q&A” section, and the “Scenarios” at the end of the chapter.
- **7–9 overall score** — Begin with the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. If you have trouble with these exercises, read the appropriate sections in “Foundation Topics.”
- **10 or more overall score** — If you want more review on these topics, skip to the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. Otherwise, move to the next chapter.

Foundation Topics

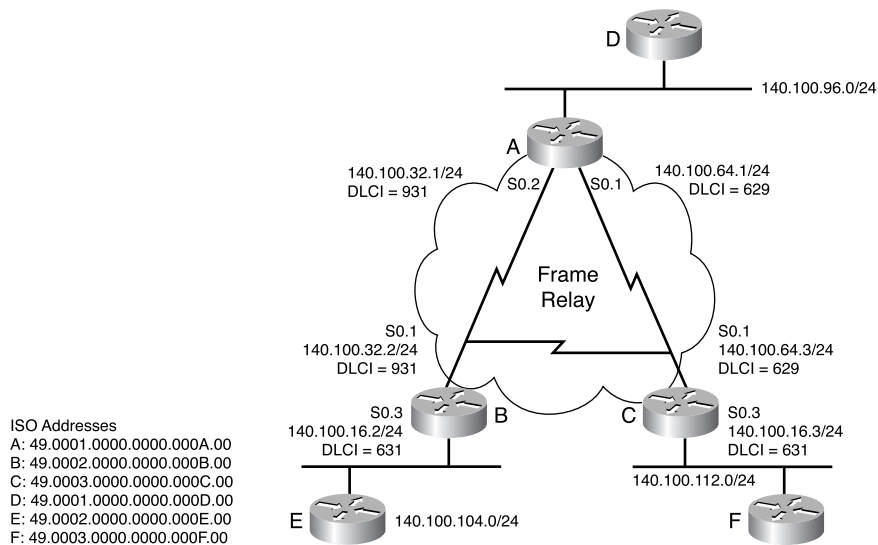
Basic Configuration of Integrated IS-IS

The preparation for configuring any routing protocol requires a thorough understanding of the network topology and a coherent addressing scheme. When you have these, the basic configuration of Integrated IS-IS is as follows:

- Step 1** Enable the routing process Integrated IS-IS with the **router isis** command.
- Step 2** Configure the Network Entity Title (NET) address, thus assigning the area with the **net network-address** router subcommand.
- Step 3** Enable Integrated IS-IS for IP on the relevant interfaces with the **ip router isis** interface subcommand.

Figure 12-1 illustrates a simple network to support the working configuration example. The figure shows the topology of the network and the addressing scheme that was used.

Figure 12-1 Basic Integrated IS-IS Configuration



Example 12-1 shows the basic configuration required to run Integrated IS-IS. The relevant commands are highlighted for easy identification, and a brief explanation of the code is inserted after an exclamation mark to make the code easier to read. Note that the router process is started and the NET address is defined immediately beneath this router process. This is done because the IS-IS address is assigned to the router and, as such, is a part of the routing process. IS-IS is started on the interface after the IS-IS router process is defined with a NET address.

Example 12-1 *Basic Integrated IS-IS Configuration*

```
Working Example for Router A
interface Ethernet0
  ip address 140.100.96.1 255.255.255.0
!assign the IP address and mask
  ip router isis
!start IS-IS on the interface
!
interface Serial0
  no ip address
  encapsulation frame-relay
  no fair-queue
  clockrate 56000
!
interface Serial0.1 point-to-point
  ip address 140.100.64.1 255.255.255.0
!assign the IP address and mask
  ip router isis
!start IS-IS on the interface
  frame-relay interface-dlci 629
!
interface Serial0.2 point-to-point
  ip address 140.100.32.1 255.255.255.0
!assign the IP address and mask
  ip router isis
!start IS-IS on the interface
  frame-relay interface-dlci 931
!
!
router isis
  net 49.0001.0000.0000.000a.00
!Start the IS-IS routing process and define the ISO address for the entire router.
!
ip classless
```

Remember that by default, Cisco routers are configured as Level 1-2 routers to avoid suboptimal routing. Also, the **cls routing** command, which turns on the routing process for the OSI protocol stack and allows it to route OSI traffic across the router, is not required for IP-only IS-IS. In some

versions of the Cisco IOS software, **cls routing** is written into the configuration file automatically by Cisco, when the routing process for IS-IS is configured.

Optional Integrated IS-IS Commands

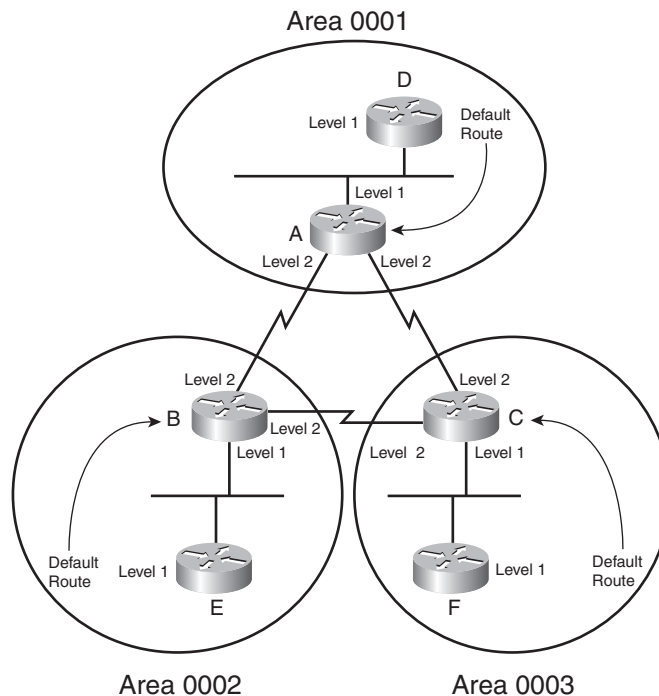
Though the commands in this section are considered optional, they might or might not be optional in your network. For instance, if you are running Frame Relay, you must carefully consider and implement a configuration of Integrated IS-IS over the nonbroadcast multiaccess (NBMA) cloud. One of many considerations is how to separate the network into areas and which routers are to take which level of responsibility.

Changing the Router from Level 1-2

The reason you change the routing level in Integrated IS-IS is to overcome the problem of Integrated IS-IS consuming resources on both the routers and the media. Using the topology in Figure 12-2, it is possible to configure Routers D, E, and F as Level 1 routers because they are internal to their areas. The following syntax shows the command structure:

```
Router(config)#ip router isis
Router(config-router)#isis circuit-type level 1
```

Figure 12-2 *Changing the Level of Routing Both for the Entire Router and at the Interface Level*



It is also possible to change the level of routing to Level 1 at the interface level by issuing the **isis circuit-type level-1** command under the interface. In Figure 12-2, this is done on the Routers A, B, and C. The Ethernet interface pointing to the stub Routers D, E, and F is configured to be running Level 1 routing, whereas the serial interface runs only Level 2 routing.

Level 1-2 routers send LSPs with an attached (ATT) bit in their Level 1 LSPs, indicating that they are attached to another area. This creates a default route in the Level 1 router pointing to the nearest Level 2 router. This allows the nearest Level 2 router to be the transit router for all data destined for another area.

As illustrated in Figure 12-2, Example 12-2 shows Router A configured as Level 1 on the Ethernet interface pointing to Router D. The other interfaces are configured as Level 2 only. The relevant commands are highlighted for easy identification.

Example 12-2 *Changing Router A Routing Level on an Interface Basis*

```
Working Example for Router A
interface Ethernet0
  ip address 140.100.96.1 255.255.255.0
  ip router isis
  isis circuit-type level-1
!Configure Level 1 routing on the interface
!
interface Serial0
  no ip address
  encapsulation frame-relay
  no fair-queue
!
interface Serial0.1 point-to-point
  ip address 140.100.64.1 255.255.255.0
  ip router isis
  frame-relay interface-dlci 629
  isis circuit-type level-2-only
! Configure Level 2 routing on the interface
!
interface Serial0.2 point-to-point
  ip address 140.100.32.1 255.255.255.0
  ip router isis
  frame-relay interface-dlci 931
  isis circuit-type level-2-only
!Configure Level 2 routing on the interface
!
!
router isis
  net 49.0001.0000.0000.000a.00
!
ip classless
```

Example 12-3 shows the configuration command to change the level of routing on Router D to be Level 1. Note that this command appears under the **router isis** command, changing the behavior of the entire process, rather than just the interface level. The relevant commands are highlighted for easy identification.

Example 12-3 *Changing the Level of Routing for the Entire Router*

```
Working Example for Router D
interface Ethernet0
 ip address 140.100.96.2 255.255.255.0
 no ip directed-broadcast
 ip router isis
!
interface Serial0
 no ip address
 no ip directed-broadcast
 no ip mroute-cache
 Shutdown
 no fair-queue
!
interface Serial0.1
 no ip directed-broadcast
!
interface Serial1
 no ip address
 no ip directed-broadcast
 Shutdown
!
!
router isis
 net 49.0001.0000.0000.000d.00
 is-type level-1
!Define the router as a Level 1 router
!
ip classless
```

Configuring Summarization

Configuring summarization is straightforward. The complexity is in the addressing scheme design. There are three rules to remember about summarizing IP routes for Integrated IS-IS:

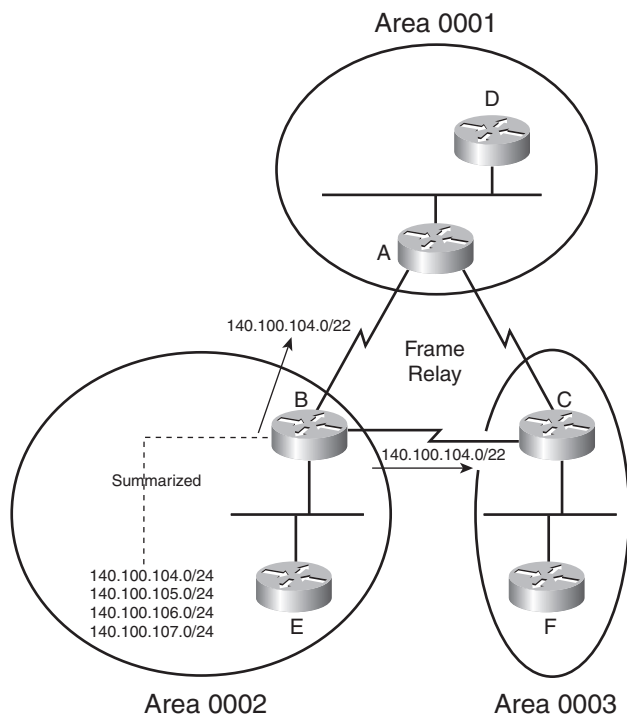
- Routes cannot be summarized within an area; that is, the protocol does not allow intra-area summarization.
- Internal routes can be summarized between areas, from a Level 1 to a Level 2 router. Thus, summarization is configured on a Level 1-2 router, which turns the routes from Level 1 routes into Level 2 routes.

- If summarization is used, all the Level 1-2 routers in the area need to be summarizing routes out of the area in the same manner. If one router is advertising the routes that are more specific (unsummarized), all the traffic will be sent to this router in accordance to the longest-match rule. This will result in suboptimal routing and might overwhelm the receiving router.

To configure summarization, enter the **summary-address** command followed by the summary address and mask under the router process. Figure 12-3 shows the topology used to support the working example shown in Example 12-4.

In Figure 12-3, the IP routes from Router B are summarized into areas 0001 and 0003. Router B is capable of summarizing routes because the router is both a Level 1 and a Level 2 router, straddling more than one area and forming the connectivity between the areas. This functionality is similar to an ABR in OSPF. Example 12-4 shows Router B summarizing routes 140.100.104.0 through 140.100.107.0 with a mask summarized from /24 to /22. The relevant command is highlighted for easy identification.

Figure 12-3 *Summarizing IP Addresses Between IS-IS Areas*



Example 12-4 *Summarization of IP Routes from Router B*

```

Working Example for Router B
interface Ethernet0
 ip address 140.100.104.1 255.255.255.0
 no ip directed-broadcast
 ip router isis
 isis circuit-type level-1
!
interface Serial0
 no ip address
 no ip directed-broadcast
 encapsulation frame-relay
 no ip mroute-cache
 no fair-queue
!
interface Serial0.2 point-to-point
 ip address 140.100.32.2 255.255.255.0
 no ip directed-broadcast
 ip router isis
 isis circuit-type level-2-only
 frame-relay interface-dlci 931
!
interface Serial0.3 point-to-point
 ip address 140.100.16.2 255.255.255.0
 no ip directed-broadcast
 ip router isis
 isis circuit-type level-2-only
 frame-relay interface-dlci 631
!
!
router isis
 summary-address 140.100.104.0 255.255.252.0
! Routes 140.100.104.0-140.100.107.0 summarized with the /22 bit mask.
 net 49.0002.0000.0000.000b.00
!
 ip classless

```

Configuring NBMA

IS-IS acknowledges only two types of network topologies: broadcast and point-to-point. If the network link is not a serial line connecting to a single router (a point-to-point network), IS-IS automatically defines the link to be broadcast. Because NBMA is neither a point-to-point nor a broadcast medium, the configuration for IS-IS over NBMA deserves a moment's consideration.

For multiaccess WAN interfaces (such as ATM, Frame Relay, and X.25), it is highly recommended that you configure the NBMA cloud as multiple point-to-point subinterfaces. This is a simpler

design that makes routing much more robust, particularly if one or more permanent virtual circuits (PVCs) in the NBMA cloud fail.

An example of an NBMA technology is Frame Relay. It is a WAN technology that is widely used and has evolved beyond the point-to-point capabilities. When configured in a full mesh, the cloud is multiaccess, although each circuit is discrete and therefore not a true broadcast medium.

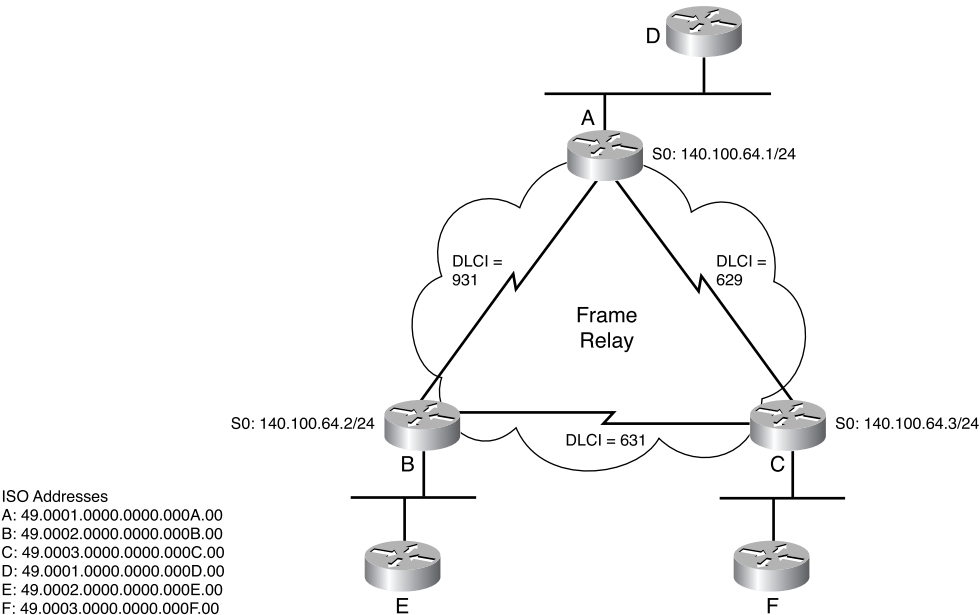
The Broadcast Configuration over NBMA

If the NBMA cloud is fully meshed, the IS-IS broadcast option is the configuration to choose. As far as IS-IS is concerned, the NBMA cloud is a broadcast medium, such as Ethernet, and so a DIS router is elected. Decisions about manually determining which router is elected should be made in reference to the topology of the network, data flow, and router capacity.

Remember that hello and routing updates are used differently than the point-to-point configuration. Therefore, you need to ensure that all the interfaces connecting into the cloud are configured in the same way; otherwise, the hellos will be rejected and no adjacency will be formed.

The configuration for IS-IS over the fully meshed Frame Relay cloud is illustrated in Figure 12-4. In Figure 12-4, the Frame Relay cloud has three fully meshed routers, which can operate as a LAN as far as IS-IS PDUs are concerned.

Figure 12-4 NBMA Frame Relay Cloud Running Broadcast Integrated IS-IS



Example 12-5 shows a working example of the configuration. In the example, the **frame-relay map ip** command maps the IP destination address to the outgoing data-link connection identifier (DLCI) and defines the interface as a broadcast interface. Integrated IS-IS uses the links as if they were truly a broadcast link and elects a DIS.

The **frame-relay map clns** command maps to the CLNS process on the destination router. Without this command, no routes appear in the IP routing table because IS-IS does not receive IS-IS frames to populate the IP routing table. Remember that these are IP routes carried in the IS-IS routing protocol. IS-IS information does not travel in IP or CLNS packets. IS-IS is encapsulated at the network layer in a frame that is similar to CLNS and those frames must be carried over Frame Relay to build the routing table.

Example 12-5 *NBMA Frame Relay Cloud Running Broadcast Integrated IS-IS*

```
Working Example for Router A
!
interface Ethernet0
 ip address 140.100.96.1 255.255.255.0
 ip router isis
 isis circuit-type level-1
!
interface Serial0
 ip address 140.100.64.1 255.255.255.0
 ip router isis
 encapsulation frame-relay
 clockrate 56000
 frame-relay map clns 629 broadcast
!Map the layer 2 DLCI address in Frame Relay to the clns process of Router C
 frame-relay map clns 931 broadcast
!Map the layer 2 DLCI address in Frame Relay to the clns process of Router B
 frame-relay map ip 140.100.64.2 931 broadcast
!Map the layer 2 DLCI address in Frame Relay to the Destination IP address of Router B
 frame-relay map ip 140.100.64.3 629 broadcast
!Map the layer 2 DLCI address in Frame Relay to the Destination IP address of Router C

 isis circuit-type level-2-only
!
interface Serial1
 No ip address
 Shutdown
!
!
router isis
 net 49.0001.0000.0000.000a.00
!
ip classless
```

The alternative solution to a broadcast configuration is to define subinterfaces and to configure each subinterface as point-to-point.

The Point-to-Point Configuration over NBMA

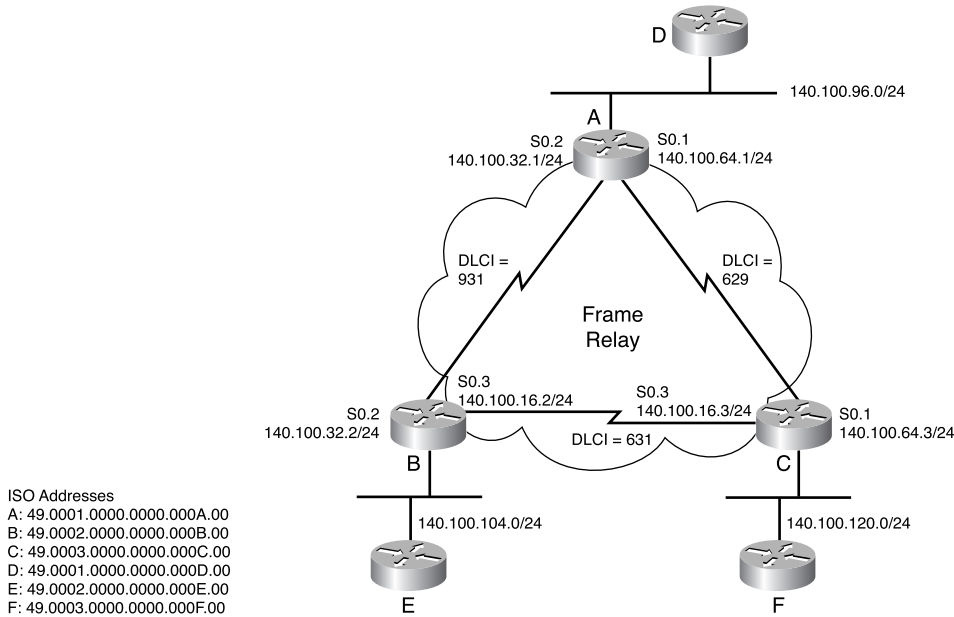
The point-to-point configuration over NBMA requires an IP subnet per link. This is the configuration suggested by Cisco for a hub and spoke topology.

The configuration is simpler, because the link is point-to-point and there is no need to configure **frame-relay map** commands. The point-to-point link is just a pipe that goes to one destination, and map commands imply a choice of destination.

As shown in the configuration in Example 12-6, it is only necessary to create subinterfaces, configure those interfaces as point-to-point, start Frame Relay, and define the DLCIs. Do not forget that in addition to configuring Frame Relay, you must start the IS-IS process for each interface.

Figure 12-5 supports the working example for this configuration. It shows the DLCI addresses and the IP addresses for the point-to-point links in addition to the ISO addresses for Routers A, B, and C.

Figure 12-5 NBMA Frame Relay Cloud Running Point-to-Point Integrated IS-IS



Example 12-6 shows the configuration for Router A to run Integrated IS-IS across the Frame Relay cloud as if it were a series of point-to-point networks.

Example 12-6 *NBMA Frame Relay Cloud Running Point-to-Point Integrated IS-IS*

```

Working Example for Router A
!
interface Ethernet0
 ip address 140.100.96.1 255.255.255.0
 ip router isis
!
interface Serial0
 no ip address
 encapsulation frame-relay
!Configure Frame Relay for the interface
 no fair-queue
 clockrate 56000
!
interface Serial0.1 point-to-point
!Configure the subinterface to be point-to-point
 ip address 140.100.64.1 255.255.255.0
 ip router isis
 frame-relay interface-dlci 629
!Configure Frame Relay and define the DLCI to the destination
interface Serial0.2 point-to-point
! Configure the subinterface to be point-to-point
 ip address 140.100.32.1 255.255.255.0
 ip router isis
 frame-relay interface-dlci 931
! Configure Frame Relay and define the DLCI to the destination
!
router isis
 net 49.0001.0000.0000.000a.00
!
 ip classless

```

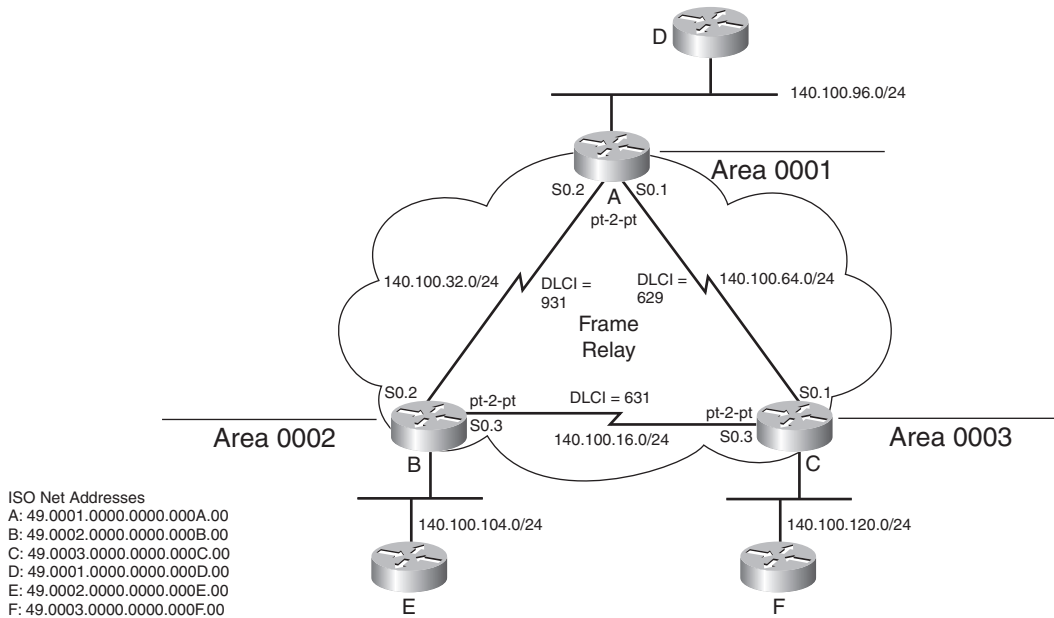
Verifying the Integrated IS-IS Operation

The ability to monitor the network enables you to optimize the network and detect problems early. Useful commands to verify the operation of Integrated IS-IS include the following **show** commands:

- **show clns neighbor**
- **show clns interface**
- **show isis database**
- **show isis database detail**

The following sections explain each of these **show** commands in more detail. The commands explained in this section correspond to the topology illustrated in Figure 12-6 and the configuration shown in Example 12-7 for Router A.

Figure 12-6 *The Network Topology for the show Commands*



Example 12-7 *The Configuration for Router A in Figure 12-6*

```

Working Example for Router A
interface Ethernet0
 ip address 140.100.96.1 255.255.255.0
 ip router isis
 isis circuit-type level-1
!
interface Serial0
 no ip address
 encapsulation frame-relay
 no fair-queue
 clockrate 56000
!
interface Serial0.1 point-to-point
 ip address 140.100.64.1 255.255.255.0
 ip router isis
 isis circuit-type level-2-only
 frame-relay interface-dlci 629
!

```

Example 12-7 *The Configuration for Router A in Figure 12-6 (Continued)*

```

interface Serial0.2 point-to-point
 ip address 140.100.32.1 255.255.255.0
 ip router isis
 isis circuit-type level-2-only
 frame-relay interface-dlci 931
!
!
router isis
 net 49.0001.0000.0000.000a.00
!
ip classless

```

The show clns neighbors Command

The **show clns neighbors** command has some of the contents of the neighbor table and the state of the link. Note that the subnetwork point of attachment (SNPA) is the MAC address of the interface. The type of routing that is used is Level 1-2.

The EXEC command has the following syntax:

```
show clns area-tag neighbors [type number] [area] [detail]
```

Table 12-2 explains the syntax of this command.

Table 12-2 *Explanation of the show clns neighbors Command*

Field	Description
<i>area-tag</i>	Used primarily in multiarea Integrated IS-IS configurations. This is a name for an Integrated IS-IS routing process on a router. Each name must be unique to the process on the router. If an area tag is specified, output is limited to the specified area.
<i>type number</i>	(Optional) Type of interface (for example, Ethernet). (Optional) Interface number (for example, Ethernet 1).
area	(Optional) Shows the CLNS multiarea adjacencies.
detail	(Optional) Shows details of each adjacency rather than a summary display.

Example 12-8 shows output for the **show clns neighbors** command.

Example 12-8 Output for the **show clns neighbors** Command

Working Example for Router A
RouterA#**show clns neighbors**

System Id	Interface	SNPA	State	Holdtime	Type	Protocol
0000.0000.000B	Se0.2	DLCI 931	Up	22	L2	IS-IS
0000.0000.000C	Se0.1	DLCI 629	Up	23	L2	IS-IS
0000.0000.000D	Et0	00e0.1e3d.d56f	Up	8	L1	IS-IS

The output of the **show clns neighbors** command shows that Router A has three neighbors. The system ID shows that the serial subinterface S0.1 has heard an LSP from 0000.0000.000C S0.1, which has the data-link DLCI Frame Relay address of 629. The protocol is IS-IS, and it is running Level 2 routing. The link is up and has 23 seconds before another Hello needs to be received. Because the Hello timer is set by default to send Hellos every 10 seconds, it should receive another Hello in 7 seconds, which will reset the holdtime. The Ethernet segment is running Level 1 routing and has a MAC address for the SNPA address.

The **show clns neighbors** command is good for quickly checking connectivity. This output shows all the neighbors—complete with the DLCI addresses and OSI system IDs—indicating that Frame Relay is correctly configured and working, as is IS-IS.

Adding the parameter **detail** to the **show clns neighbors** command gives information about each neighbor and the connection to that neighbor. Example 12-9 shows output for the **show clns neighbors detail** command. Here the area address for the neighbor, the IP address of the transmitting interface, and the length of time that the interface has been up are shown. This command gives information that enables you to verify the addressing scheme of the network.

Example 12-9 Output for the **show clns neighbors detail** Command

Working Example for Router A
RouterA#**show clns neighbor detail**

System Id	Interface	SNPA	State	Holdtime	Type	Protocol
0000.0000.000B	Se0.2	DLCI 931	Up	27	L2	IS-IS
Area Address(es): 49.0002						
IP Address(es): 140.100.32.2*						
Uptime: 00:05:17						
0000.0000.000C	Se0.1	DLCI 629	Up	28	L2	IS-IS
Area Address(es): 49.0003						
IP Address(es): 140.100.64.3*						
Uptime: 00:05:22						
0000.0000.000D	Et0	00e0.1e3d.d56f	Up	7	L1	IS-IS

Example 12-9 *Output for the show clns neighbors detail Command (Continued)*

```
Area Address(es): 49.0001
IP Address(es): 140.100.96.2*
Uptime: 00:15:01
```

Table 12-3 explains the meaning of the fields in the output. The command explained is with the **detail** parameter to cover all the fields shown in both output examples.

Table 12-3 *Explanation of the show clns neighbors detail Command Output*

Field	Description
System Id	The system address. It identifies a system in an area. Cisco defines this to be a 6-octet identifier.
SNPA	Subnetwork point of attachment. This is the data-link address.
Interface	Interface through which the neighbor was learned.
State	State of the ES or IS. The states are as follows: • Init—The status of the adjacency is initializing. The router is waiting for an Integrated IS-IS Hello message. Until the Hello is received, the neighbor is not adjacent. • Up—Considers the other system reachable. • Down—Considers the other system unreachable.
Holdtime	Number of seconds before this adjacency entry times out.
Type	The types of adjacency are as follows: • ES—End-system adjacency. This could be discovered via the ES-IS protocol or statically configured. • Router adjacency—This could be discovered via the ES-IS protocol or statically configured.
L1	(Optional) Displays the router adjacencies for Level 1.
L1L2	(Optional) Displays the router adjacencies for Level 1 and 2.
L2	(Optional) Displays the router adjacencies for Level 2.
Protocol	The routing protocol used to learn the adjacency. The protocol might be ES-IS, IS-IS, ISO IGRP, Static, or DECnet.

The show clns interface Command

The misconfiguration of the interface for Integrated IS-IS results in the inability to create adjacencies. Typically, the error is a simple mismatch of parameters, which can be seen by using the **show clns interface** command.

The neighbor database tells you of one neighbor, and the interface indicates one adjacency. Because this is a LAN interface, it is possible to identify the DIS. The circuit ID shows the pseudonode ID, which has a value greater than 0x00 in the octet after the system ID. Note that a DIS is elected for both Level 1 and Level 2 routing. Because no priority has been manually configured, the tiebreaker used to elect the DIS is the highest SNPA on the segment. In this case, the SNPA is the MAC address. Note that the default metric is 10 and the priority is 64.

The EXEC command has the following syntax:

```
Router#show clns interface [type number]
```

In this example, Frame Relay is configured with point-to-point links. Because there is only one other router on this link, there is no need for a DIS to be elected. Therefore, the circuit ID shows the system ID of a router, rather than a pseudonode. Remember, the octet following the system ID indicates whether this ID is a pseudonode representing the multiaccess link. If the ID is that of a pseudonode, the system ID is that of the DIS, with the next octet showing a nonzero value such as 0x01.

This makes more sense when you look at the Ethernet interface. The Ethernet 0 interface has the Level 1 circuit ID as A.01. This indicates that the DIS for Level 1 is Router A. The octet following the circuit ID of 01 has a nonzero value, indicating the ID of a pseudonode.

Because this is a Level 1-2 router, there is also a circuit ID for the Level 2 adjacency. This is relevant only on the Ethernet interface because it is the only multiaccess link. Note that the value for the Level 2 DIS is that of Router A. Router D has been configured as a Level 1 router and, as such, cannot communicate Level 2 updates.

Example 12-10 shows output for the **show clns interface** command.

Example 12-10 Output for the show clns interface Command

```
Working Example for Router A
RouterA#show clns interface
Ethernet0 is up, line protocol is up
  Checksums enabled, MTU 1497, Encapsulation SAP
  ERPDUs enabled, min. interval 10 msec.
  CLNS fast switching enabled
  CLNS SSE switching disabled
  DEC compatibility mode OFF for this interface
  Next ESH/ISH in 47 seconds
```


Example 12-10 *Output for the show clns interface Command (Continued)*

```

Routing Protocol: IS-IS
  Circuit Type: level-1-2
  Interface number 0x0, local circuit ID 0x1
  Level-1 Metric: 10, Priority: 64, Circuit ID: A.01
  Number of active level-1 adjacencies: 0
  Level-2 Metric: 10, Priority: 64, Circuit ID A.01
  Number of active level-2 adjacencies: 1
  Next IS-IS LAN Level-1 Hello in 1 seconds
  Next IS-IS LAN Level-2 Hello in 1 seconds
Serial0 is up, line protocol is up
  CLNS protocol processing disabled
Serial0.1 is up, line protocol is up
  Checksums enabled, MTU 1500, Encapsulation FRAME-RELAY
  ERPDUs enabled, min. interval 10 msec.
  CLNS fast switching disabled
  CLNS SSE switching disabled
  DEC compatibility mode OFF for this interface
  Next ESH/ISH in 58 seconds
Routing Protocol: IS-IS
  Circuit Type: level-1-2
  Interface number 0x1, local circuit ID 0x100
  Level-1 Metric: 10, Priority: 64, Circuit ID: A.00
  Number of active level-1 adjacencies: 0
  Level-2 Metric: 10, Priority: 64, Circuit ID: A.00
  Number of active level-2 adjacencies: 1
  Next IS-IS Hello in 2 seconds
Serial0.2 is up, line protocol is up
  Checksums enabled, MTU 1500, Encapsulation FRAME-RELAY
  ERPDUs enabled, min. interval 10 msec.
  CLNS fast switching disabled
  CLNS SSE switching disabled
  DEC compatibility mode OFF for this interface
  Next ESH/ISH in 24 seconds
Routing Protocol: IS-IS
  Circuit Type: level-1-2
  Interface number 0x2, local circuit ID 0x101
  Level-1 Metric: 10, Priority: 64, Circuit ID: A.00
  Number of active level-1 adjacencies: 0
  Level-2 Metric: 10, Priority: 64, Circuit ID: A.00
  Number of active level-2 adjacencies: 1
  Next IS-IS Hello in 886 milliseconds

```

Table 12-4 explains the meaning of the fields in the output screen.

Table 12-4 *Explanation of the show clns interface Command Output*

Field	Description
Checksum	This may be either enabled or disabled.
MTU	The number following MTU is the maximum transmission size for a packet on this interface. Note that the MTU is not 1500 because 3 bytes are taken by the OSI header.
Encapsulation	The encapsulation is always SAP (ISO1).
Routing Protocol	This indicates whether ES-IS or IS-IS is running. This determines the type of Hellos sent.
Circuit Type	This indicates whether the link is enabled for L1, L2, or L1-2 routing. This determines the Hellos that are sent.
Level-1 Metric	The metric value for the outbound interface for Level 1 routing. Note that the default setting is 10.
Priority	The priority setting for DIS election. Note that the default of 64 is used.
Circuit ID	This identifies the DIS for L1 if there is one on the link.
Number of Active Level-1 Adjacencies	The number of L1 adjacencies formed on this link.
Level-2 Metric	The metric value for the outbound interface for Level 2 routing. Note that the default setting is 10.
Priority	The priority setting for DIS election. Note that the default of 64 is used.
Circuit ID	This identifies the DIS for L2 if there is one on the link.
Number of Active Level-2 Adjacencies	The number of L2 adjacencies formed on this link.
Next IS-IS LAN Level-1 Hello	The number of seconds before the next Hello is expected. This is a good troubleshooting tool, because you can repeat the command to see whether the Hellos are received.
Next IS-IS LAN Level-2 Hello	The number of seconds before the next Hello is expected. This is a good troubleshooting tool, because you can repeat the command to see whether the Hellos are received.

The show isis database Command

The **show isis database** command shows the LSPs held in the local database. The LSP ID shows the system ID of the generating router and whether this LSP is from a router or a pseudonode. The

last octet shows whether the LSP was too large to fit into one PDU; a nonzero value in this field indicates that this is a fragment of an LSP. The number states the fragment number.

Because IS-IS is a link-state protocol, the database should be identical on every router of the same level within the area.

The EXEC command has the following syntax:

```
show isis area-tag database [level-1] [level-2] [l1] [l2] [detail] [lspid]
```

Table 12-5 explains the syntax of this command.

Table 12-5 Explanation of the **show isis database** Command

Field	Description
<i>area-tag</i>	Used primarily in multiarea Integrated IS-IS configurations. This is a name for an Integrated IS-IS routing process on a router. Each name must be unique to the process on the router. If an area tag is specified, output is limited to the specified area. Required for multiarea Integrated IS-IS configuration. Optional for conventional Integrated IS-IS configuration.
level-1	(Optional) Displays the Integrated IS-IS link-state database for Level 1.
level-2	(Optional) Displays the Integrated IS-IS link-state database for Level 2.
l1	(Optional) Abbreviation for level-1 .
l2	(Optional) Abbreviation for level-2 .
detail	(Optional) Shows the contents of each LSP rather than a summary display.
lspid	(Optional) Identifier for the link-state PDU. Shows the contents of the specified individual LSP.

Example 12-11 shows output for the **show isis database** command.

Example 12-11 Output for the **show isis database** Command

RouterA# show isis database				
IS-IS Level-1 Link State Database:				
LSPID	LSP Seq Num	LSP Checksum	LSP Holddtime	ATT/P/OL
A.00-00	* 0x00000017	0x76D5	876	1/0/0
IS-IS Level-2 Link State Database:				
LSPID	LSP Seq Num	LSP Checksum	LSP Holddtime	ATT/P/OL
A.00-00	* 0x00000018	0xB74F	881	0/0/0
0000.0000.000B.00-00	0x0000001A	0xB561	872	0/0/0
0000.0000.000B.01-00	0x00000016	0x6045	1095	0/0/0

continues

Example 12-11 Output for the **show isis database** Command (Continued)

C.00-00	0x0000001E	0x6267	869	0/0/0
C.01-00	0x00000002	0xF25F	958	0/0/0
0000.0000.000E.00-00	0x00000018	0x010A	858	0/0/0
0000.0000.000D.00-00	0x0000001A	0x413C	985	0/0/0
0000.0000.000D.04-00	0x00000017	0xFCA0	1006	0/0/0

Table 12-6 explains the meaning of the fields in the output. Note that both Level 1 and Level 2 databases are shown because the router is running the default configuration of Level 1-2 routing. All the LSP information is contained in these databases, including the LSPs generated by the router itself. An asterisk marks these entries in the output.

Table 12-6 Explanation of the **show isis database** Command

Field	Description
LSPID	The LSPID indicates the transmitting router. The system ID is followed by two octets. If the first octet has a value greater than 0x00, this indicates that the ID is that of the pseudonode and the system ID is that of the DIS. The last octet is the Fragment bit. If the value is 0x00, the entire LSP was carried in one LSP. If the value is greater than 0x00, it indicates that this PDU is the <i>n</i> th fragment of the LSP.
LSP Seq Num	Sequence number for the LSP. This allows routers to determine whether they have received the latest information from the source and ensures the integrity of the link-state database.
LSP Checksum	Checksum of the entire LSP packet.
LSP Holdtime	Amount of time the LSP remains valid (in seconds). An LSP holdtime of 0 indicates that this LSP was purged and is being removed from the link-state database.
ATT	The LSP indicates that this router is a Level 2 router with a path out of the area. Level 1 routers use the presence of this bit to identify the closest Level 1-2 router to send their out-of-area data.
P	Partition repair capability. Not supported by Cisco.
OL	The Overload bit indicates that the router has an incomplete database because of memory overload and is therefore not used for transit data.

The **show isis database detail** Command

The **show isis database detail** command shows the complete LSP and the values for the individual fields. The EXEC command has the same syntax structure as shown in the **show isis database** command in the preceding section.

Example 12-12 shows output for the **show isis database detail** command. The highlighted lines show that there are two databases: one for the Level 1 routing and the other for the Level 2 routing. The example shows, for each LSP, the area and the IP address of the transmitting interface and the metric cost to the IP routes it knows. The default metric is a cost of 10; therefore, a metric of 20 indicates a route that is two hops away.

Example 12-12 *Output for the show isis database detail Command*

```
RouterA#show isis database detail
```

IS-IS Level-1 Link State Database:				
LSPID	LSP Seq Num	LSP Checksum	LSP Holdtime	ATT/P/OL
A.00-00	* 0x00000017	0x76D5	873	1/0/0
Area Address: 49.0001				
NLPID: 0xCC				
Hostname:A				
IP Address: 140.100.32.1				
Metric: 10	IP 140.100.96.0 255.255.255.0			
Metric: 10	IP 140.100.64.0 255.255.255.0			
Metric: 10	IP 140.100.32.0 255.255.255.0			
Metric: 10	IS A.01			
IS-IS Level-2 Link State Database:				
LSPID	LSP Seq Num	LSP Checksum	LSP Holdtime	ATT/P/OL
A.00-00	* 0x00000018	0xB74F	877	0/0/0
Area Address: 49.0001				
NLPID: 0xCC				
Hostname: A				
IP Address: 140.100.32.1				
Metric: 10	IS 0000.0000.000B.00			
Metric: 10	IS C.00			
Metric: 10	IS 0000.0000.000D.04			
Metric: 10	IP 140.100.96.0 255.255.255.0			
Metric: 10	IP 140.100.64.0 255.255.255.0			
Metric: 10	IP 140.100.32.0 255.255.255.0			
0000.0000.000B.00-00	0x0000001A	0xB561	868	0/0/0
Area Address: 49.0002				
NLPID: 0xCC				
IP Address: 140.100.16.2				
Metric: 10	IS 0000.0000.000B.01			
Metric: 10	IS A.00			
Metric: 10	IS C.00			
Metric: 10	IP 140.100.104.0 255.255.255.0			
Metric: 20	IP 140.100.105.0 255.255.255.0			
Metric: 20	IP 140.100.106.0 255.255.255.0			
Metric: 20	IP 140.100.107.0 255.255.255.0			
Metric: 10	IP 140.100.32.0 255.255.255.0			
Metric: 10	IP 140.100.16.0 255.255.255.0			

continues

Example 12-12 *Output for the show isis database detail Command (Continued)*

0000.0000.000B.01-00	0x00000016	0x6045	1089	0/0/0
Metric: 0	IS 0000.0000.000B.00			
Metric: 0	IS 0000.0000.000E.00			
C.00-00	0x0000001E	0x6267	863	0/0/0
Area Address: 49.0003				
NLPID: 0xCC				
Hostname: C				
IP Address: 140.100.100.1				
Metric: 10	IS C.02			
Metric: 10	IS A.00			
Metric: 10	IS 0000.0000.000B.00			
Metric: 10	IP 140.100.100.0 255.255.255.0			
Metric: 10	IP 140.100.64.0 255.255.255.0			
Metric: 10	IP 140.100.16.0 255.255.255.0			
C.01-00	0x00000002	0xF25F	951	0/0/0
Metric: 0	IS C.00			
0x00000018	0x010A	850	0/0/0	
Area Address: 49.0002				
0000.0000.000E.00-00	NLPID: 0xCC			
IP Address: 140.100.105.1				
Metric: 10	IS 0000.0000.000B.01			
Metric: 10	IP 140.100.104.0 255.255.255.0			
Metric: 10	IP 140.100.105.0 255.255.255.0			
Metric: 10	IP 140.100.106.0 255.255.255.0			
Metric: 10	IP 140.100.107.0 255.255.255.0			
Metric: 20	IP 140.100.32.0 255.255.255.0			
Metric: 20	IP 140.100.16.0 255.255.255.0			
0000.0000.000D.00-00	0x0000001A	0x413C	976	0/0/0
Area Address: 49.0003				
NLPID: 0xCC				
IP Address: 140.100.97.1				
Metric: 10	IS 0000.0000.000D.04			
Metric: 10	IP 140.100.96.0 255.255.255.0			
Metric: 10	IP 140.100.97.0 255.255.255.0			
Metric: 10	IP 140.100.98.0 255.255.255.0			
Metric: 10	IP 140.100.99.0 255.255.255.0			
0000.0000.000D.04-00	0x00000017	0xFCA0	996	0/0/0
Metric: 0	IS 0000.0000.000D.00			
Metric: 0	IS A.00			

Table 12-7 explains the meaning of the fields in the output.

Table 12-7 *Explanation of the **show isis database detail** Command*

Field	Description
Area Address	Area addresses that may be reached from this router. L1 LSPs describe the area addresses configured manually on the originating router. L2 LSPs describe the area addresses for the area to which this route belongs.
Metric	The cost of the outbound interface between the originating router and the advertised neighbor, or the metric of the compound cost between the advertising router and the advertised destination.

Troubleshooting the Integrated IS-IS Operation

Unfortunately, even after the most careful planning, configurations can fail to work, and the most scrutinized networks can break. The **show** commands are essential for hunting down problems. Careful documentation of the troubleshooting steps taken enable the administrator to build a solid understanding of the problem and systematically eliminate possible problem sources. Having this log is invaluable when, if all else fails, you have to elicit the help of the Technical Assistance Center (TAC) engineers at Cisco. The following commands are useful for troubleshooting Integrated IS-IS:

- **show isis spf-log**
- **debug** commands

The following sections describe these commands in greater detail.

The **show isis spf-log** Command

The **show isis spf-log** command explains a great deal about the SPF calculations on the router. It gives the events that triggered SPF for the last 20 occurrences.

To display how often and why the router has run a full SPF calculation, use the **show isis spf-log EXEC** command.

Example 12-13 shows output for the **show isis spf-log** command. The highlighted lines show the normal periodic SPF calculations. The example also shows that Router A sent an LSP because it had set the ATT bit, and later it generated a new LSP when new adjacencies came online. This output shows the healthy workings of an IS-IS network, but the command is very useful if the network appears unstable, with routes appearing and disappearing.

Example 12-13
Output for the show isis spf-log Command

RouterA#show isis spf-log						
Level 1 SPF log						
When	Duration	Nodes	Count	Last trigger LSP	Triggers	
04:23:24	0	1	1	A.00-00	TLVCODE	
04:08:46	0	1	1		PERIODIC	
03:53:46	0	1	1		PERIODIC	
03:38:46	0	1	1		PERIODIC	
03:23:46	0	1	1		PERIODIC	
03:08:46	0	1	1		PERIODIC	
02:53:46	0	1	1		PERIODIC	
02:38:46	0	1	1		PERIODIC	
02:23:46	0	1	1		PERIODIC	
02:08:46	0	1	1		PERIODIC	
01:53:46	0	1	1		PERIODIC	
01:38:46	0	1	1		PERIODIC	
01:23:46	0	1	1		PERIODIC	
01:08:46	0	1	1		PERIODIC	
00:53:46	0	1	1		PERIODIC	
00:38:47	0	1	1		PERIODIC	
00:23:47	0	1	1		PERIODIC	
00:15:14	0	1	2	A.00-00	ATTACHFLAG LSPHEADER	
00:08:46	0	1	1		PERIODIC	
Level 2 SPF log						
When	Duration	Nodes	Count	First trigger LSP	Triggers	
03:53:48	0	1	1		PERIODIC	
03:38:48	0	1	1		PERIODIC	
03:23:48	0	1	1		PERIODIC	
03:08:48	0	1	1		PERIODIC	
02:53:48	0	1	1		PERIODIC	
02:38:48	0	1	1		PERIODIC	
02:23:48	0	1	1		PERIODIC	
02:08:48	0	1	1		PERIODIC	
01:53:48	0	1	1		PERIODIC	
01:38:48	0	1	1		PERIODIC	
01:23:48	0	1	1		PERIODIC	
01:08:48	0	1	1		PERIODIC	
00:53:48	0	1	1		PERIODIC	
00:38:48	0	1	1		PERIODIC	
00:23:48	0	1	1		PERIODIC	
00:15:22	0	3	5	A.00-00	NEWADJ LSPHEADER TLVCONTENT	
00:08:48	0	3	1		PERIODIC	
00:05:44	4	4	4	A.00-00	NEWADJ TLVCONTENT	
00:05:38	4	7	5	0000.0000.000B.00-00	LSPHEADER TLVCONTENT	

Table 12-8 explains the meaning of the fields in the output screen.

Table 12-8 *Explanation of the `show isis spf-log` Command*

Field	Description
When	The amount of time since a full SPF calculation occurred (hh:mm:ss). The last 19 occurrences are shown.
Duration	Number of milliseconds it took to complete this SPF run. This is elapsed time as opposed to CPU time.
Nodes	Number of routers and pseudonodes (LANs) calculated in the SPF run.
Count	This count shows the number of events (such as receiving new LSPs) that occurred while the router was waiting before running full SPF. The router waits five seconds after the first triggered LSP to ensure that all the information is received.
Last Trigger LSP	Whenever a full SPF calculation is triggered by the arrival of a new LSP, the ID of the LSP is stored. The LSPID can indicate where to start troubleshooting for any routing instability in an area. If multiple LSPs are causing an SPF run, only the LSPID of the last received LSP is remembered.
Trigger	A list of all events that triggered a full SPF calculation.

The debug Commands

The **debug** command is a helpful troubleshooting tool, but it does have certain disadvantages of which you must be aware. The **debug** command has the highest process priority and forces the router into process switching. It is capable of consuming all the resources on the router, thus becoming the problem instead of helping to solve the problem. It is important to turn on **debug** just for the specific task to be monitored and to turn it off as soon as the data has been gathered. The **no** form of this command disables debugging output. You should direct the output to a log file, because each character sent to the console forces processor interrupt; in this way, the data can be perused with care.

The **debug** commands are EXEC commands. Table 12-9 lists the options available for monitoring Integrated IS-IS.

Table 12-9 *The debug Command Options for Integrated IS-IS*

Command Option	Description
debug isis adjacencies-packets	Displays information on all adjacency-related activity. This includes the following: • Hello packets (sent and received) • Any changes in the state of an adjacency in Integrated IS-IS
debug isis spf-statistics	Displays statistical information about building routes between routers. Using the statistical information provided, one can determine how long it takes to place a Level 1 IS or Level 2 IS on the shortest path tree (SPT) using the IS-IS protocol.
debug isis update-packets	Displays SNPs (CSNPs and PSNPs) and LSPs that are detected by the router.

Foundation Summary

The “Foundation Summary” section of each chapter lists the most important facts from the chapter. Although this section does not list every fact from the chapter that will be on your exam, a well-prepared candidate should, at a minimum, know all the details in each “Foundation Summary” before going to take the exam.

Table 12-10 summarizes the commands covered in this chapter.

Table 12-10 *Summary of Commands*

Command	Function	Location of Command
Router# router isis	Starts the Integrated IS-IS process on the router	An EXEC command in configuration mode
Router(config-router)# net network-address	Defines the NET address of the router	A subrouter command, placed under router isis
Router(config-router)# is-type level-1	Defines whether the router is acting as an L1, L2, or L1-2 router	A subrouter command, placed under router isis
Router(config-router)# summary-address address mask	Enables you to summarize the networks manually	A subrouter command, placed under router isis
Router(config-if)# ip router isis	Turns on Integrated IS-IS routing for the interface	An interface command, placed under interface [type number]
Router(config-if)# clsns router isis	Turns CLNS routing at the interface to allow the L2 LSPs to be unraveled so that the ATT bit can be seen	An interface command, placed under interface [type number]
Router(config-if)# frame-relay map ip destination address DLCI broadcast	Maps the Frame Relay address to the destination and defines the link as a broadcast link to be treated as a LAN	An interface command, placed under interface [type number]
Router# show clsns neighbor detail	Displays information drawn from the neighbor table	An EXEC command
Router# show clsns interface	Displays information about Integrated IS-IS Hello and timers seen on the interface	An EXEC command
Router# show isis database	Shows the link-state database	An EXEC command
Router# show isis spf-log	Show why and how often the SPF algorithm was calculated (listing the last 20 occurrences)	An EXEC command

Table 12-11 shows the **debug** command options discussed in this chapter.

Table 12-11 *The debug Command Options for Integrated IS-IS*

Command Option	Description
debug isis adjacencies-packets	Displays information on all adjacency-related activity. This includes the following: • Hello packets (sent and received) • Any changes in the state of an adjacency in Integrated IS-IS
debug isis spf-statistics	Displays statistical information about building routes between routers. Using the statistical information provided, one can determine how long it takes to place a Level 1 IS or Level 2 IS on the shortest path tree (SPT) using the IS-IS protocol.
debug isis update-packets	Displays SNPs (CSNPs and PSNPs) and LSPs that are detected by the router.

Q&A

As mentioned in the introduction, “All About the CCNP, CCDP, and CCIP Certifications,” you have two choices for review questions. The questions that follow next give you a bigger challenge than the exam itself by using an open-ended question format. By reviewing now with this more difficult question format, you can exercise your memory better and prove your conceptual and factual knowledge of this chapter. The answers to these questions are found in Appendix A.

For more practice with examlike question formats, including questions using a router simulator and multichoice questions, use the exam engine on the CD-ROM.

1. What command is used to configure the Integrated IS-IS router process?
2. What is the default routing level on a Cisco router?
3. What command is used to configure Integrated IS-IS routing on the interface?
4. How is the NET address configured on the router?
5. What command is used to show the state of adjacencies on the router?
6. What command identifies the designated intermediate system router for your LAN?
7. Explain briefly what **show isis database** reveals.
8. What command reveals the trigger for the last SPF calculation on the router?
9. For Frame Relay, when would you configure the **map** command with the broadcast parameter?
10. Which command is used to display all update packets that are both received and sent by a router?
11. State two reasons why a router may not be able to find a neighbor.
12. Which command shows the LSPs in detail?
13. How would you ensure that an adjacency has been established?
14. What are the steps required for a basic configuration in IS-IS?
15. Give the commands required to summarize the networks 10.10.0.0 through to 10.10.255.0 into another area of IS-IS.
16. Which command verifies the circuit type and the metric?
17. Which network topology defaults in the Broadcast mode?
18. For which WAN topology is a point-to-point configuration recommended?
19. When is it necessary to map CLNS to the DLCI?
20. In the **show clns interface** command, it is possible to identify the DIS on a multiaccess link. Which field in the output screen of this command would show the DIS for the segment?

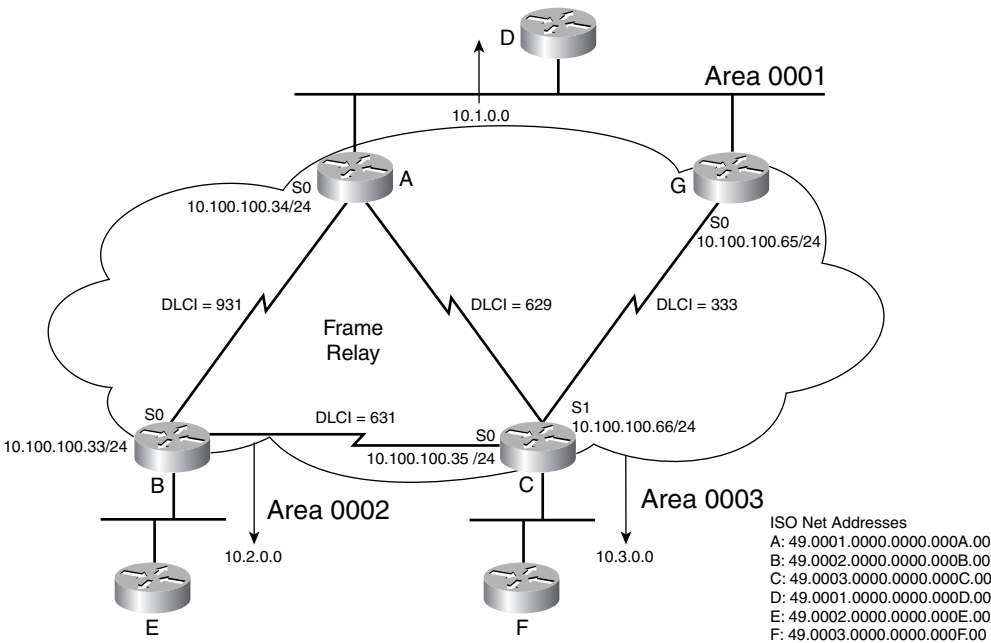
Scenarios

The following scenarios and questions are designed to draw together the content of the chapter and to exercise your understanding of the concepts. There is not necessarily a right answer. The thought process and practice in manipulating the concepts are the goals of this section. The answers to the scenario questions are found at the end of this chapter.

Scenario 12-1

The ISP Flying Data has recently converted from OSPF to Integrated IS-IS. The migration was relatively painless. The company uses the private network 10.0.0.0 for IP and the private ISO addressing with AFI 49. They created a hierarchical addressing structure. See Figure 12-7 to see this addressing scheme.

Figure 12-7 Diagram for Scenario 12-1



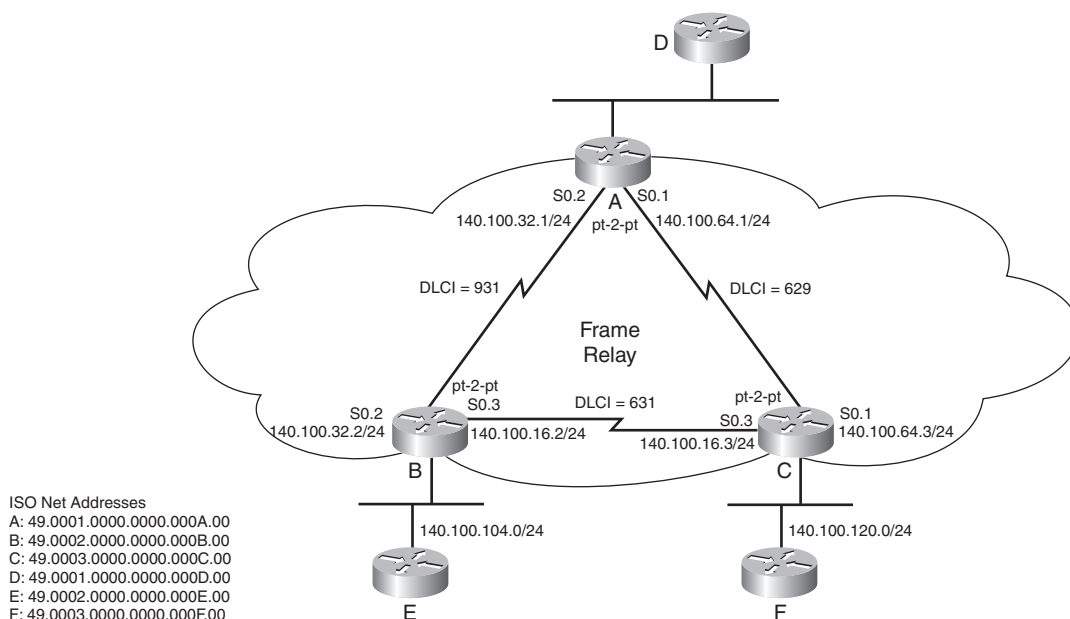
The addressing of the network was a large project, with all the pitfalls that accompany such a major exercise. The network is now stable, and it is time to configure the WAN connections using multipoint Frame Relay.

1. Issue the commands that will allow Router A to use Integrated IS-IS routing across the NBMA cloud as if the cloud were a broadcast medium. Refer to Figure 12-7 for the addressing scheme.
2. The WAN is a Frame Relay cloud, and Router G has a point-to-point link with Router C. Issue the commands for Router C that configure the link for Integrated IS-IS as a point-to-point link.
3. To reduce bandwidth consumption and to hide some network detail, summarization has been suggested as a solution over the WAN links. Issue the commands for Router A that will summarize the networks behind this router with a prefix of /16 across the WAN.

Scenario 12-2

Given the configuration of Integrated IS-IS in Example 12-14 and the output screen in Example 12-15, perform the tasks and answer the questions listed. The WAN has light user traffic and has a fully meshed configuration, as shown in Figure 12-8.

Figure 12-8 Diagram for Scenario 12-2



Example 12-14 *Router B's Configuration File*

```
RouterB#show running config
Building configuration...

Current configuration:
!
Version 12.0
Service timestamps debug uptime
Service timestamps log uptime
no service password-encryption
!
hostname B
!
!
ip subnet-zero
ip tcp synwait-time 5
no ip domain-lookup
clsns routing
!
!
Interface Ethernet0
 ip address 140.100.104.1 255.255.255.0
 no ip directed-broadcast
 ip router isis
!
Interface Serial0
 no ip address
 no ip directed-broadcast
 encapsulation frame-relay
 no ip mroute-cache
 no fair-queue
 clockrate 56000
!
Interface Serial0.2 point-to-point
 ip address 140.100.32.2 255.255.255.0
 no ip directed-broadcast
 ip router isis
 frame-relay interface-dlci 931
!
Interface Serial0.3 point-to-point
 ip address 140.100.16.2 255.255.255.0
 no ip directed-broadcast
 ip router isis
 frame-relay interface-dlci 631
!
Interface Serial1
 no ip address
```


Example 12-14 *Router B's Configuration File (Continued)*

```

no ip directed-broadcast
Shutdown
!
router isis
net 49.0002.0000.0000.000b.00
!
ip classless
!
!
alias exec c conf t
alias exec s sh run
alias exec ii sh ip int brie
!
line con 0
exec-timeout 30 0
privilege level 15
logging synchronous
transport input none
line aux 0
line vty 0 4
!
End

```

Example 12-15 *The show clns interface Command from Router B*

```

B# show clns interface
Ethernet0 is up, line protocol is up
  Checksums enabled, MTU 1497, Encapsulation SAP
  ERPDUs enabled, min. interval 10 msec.
  RDPDUs enabled, min. interval 100 msec., Addr Mask enabled
  Congestion Experienced bit set at 4 packets
  CLNS fast switching enabled
  CLNS SSE switching disabled
  DEC compatibility mode OFF for this interface
  Next ESH/ISH in 15 seconds
  Routing Protocol: IS-IS
    Circuit Type: level-1-2
    Interface number 0x0, local circuit ID 0x1
    Level-1 Metric: 10, Priority: 64, Circuit ID: 0000.0000.000B.01
    Number of active level-1 adjacencies: 1
    Level-2 Metric: 10, Priority: 64, Circuit ID: 0000.0000.000B.01
    Number of active level-2 adjacencies: 1
    Next IS-IS LAN Level-1 Hello in 678 milliseconds
    Next IS-IS LAN Level-2 Hello in 1 seconds

```

continues

Example 12-15 *The show clns interface Command from Router B (Continued)*

```

Serial0 is up, line protocol is up
  CLNS protocol processing disabled
Serial0.1 is up, line protocol is up
  CLNS protocol processing disabled
Serial0.2 is up, line protocol is up
  Checksums enabled, MTU 1500, Encapsulation FRAME-RELAY
  ERPDUs enabled, min. interval 10 msec.
  RDPDUs enabled, min. interval 100 msec., Addr Mask enabled
  Congestion Experienced bit set at 4 packets
  CLNS fast switching disabled
  CLNS SSE switching disabled
  DEC compatibility mode OFF for this interface
  Next ESH/ISH in 43 seconds
  Routing Protocol: IS-IS
    Circuit Type: level-1-2
    Interface number 0x2, local circuit ID 0x101
    Level-1 Metric: 10, Priority: 64, Circuit ID: 0000.0000.000B.01
    Number of active level-1 adjacencies: 0
    Level-2 Metric: 10, Priority: 64, Circuit ID: 0000.0000.000B.01
    Number of active level-2 adjacencies: 1
    Next IS-IS Hello in 2 seconds
Serial0.3 is up, line protocol is up
  Checksums enabled, MTU 1500, Encapsulation FRAME-RELAY
  ERPDUs enabled, min. interval 10 msec.
  RDPDUs enabled, min. interval 100 msec., Addr Mask enabled
  Congestion Experienced bit set at 4 packets
  CLNS fast switching disabled
  CLNS SSE switching disabled
  DEC compatibility mode OFF for this interface
  Next ESH/ISH in 11 seconds
  Routing Protocol: IS-IS
    Circuit Type: level-1-2
    Interface number 0x1, local circuit ID 0x100
    Level-1 Metric: 10, Priority: 64, Circuit ID: 0000.0000.000C.01
    Number of active level-1 adjacencies: 0
    Level-2 Metric: 10, Priority: 64, Circuit ID: 0000.0000.000B.00
    Number of active level-2 adjacencies: 1
    Next IS-IS Hello in 3 seconds

```

1. Identify the DIS on the Ethernet segment. How was this information apparent?
2. If Router A died, what would the effect be on the network?
3. Is summarization possible only on the routers entering the WAN cloud, or is it possible on the networks not shown in the figure, but on the other side of the routers? Give reasons for your answers.

Scenario Answers

The answers provided in this section are not necessarily the only possible answers to the questions. The questions are designed to test your knowledge and to give practical exercise in certain key areas. This section is intended to test and exercise skills and concepts detailed in the body of this chapter.

If your answer is different, ask yourself whether it follows the tenets explained in the answers provided. Your answer is correct not if it matches the solution provided in the book, but rather if it has included the principles of design laid out in the chapter.

In this way, the testing provided in these scenarios is deeper: It examines not only your knowledge, but also your understanding and ability to apply that knowledge to problems.

If you do not get the correct answer, refer back to the text and review the subject tested. Be certain to also review your notes on the question to ensure that you understand the principles of the subject.

Scenario 12-1 Answers

1. *Issue the commands that will allow Router A to use Integrated IS-IS routing across the NBMA cloud as if the cloud were a broadcast medium. Refer to Figure 12-7 for the addressing scheme.*

The highlighted portion of the configuration file in Example 12-16 shows the configuration of IS-IS across the Frame Relay cloud, using the broadcast technology and LAN Hellos. The **frame-relay map ip** command maps the IP destination address to the outgoing DLCI and defines the interface as a broadcast interface.

The **frame-relay map clns** command maps to the CLNS process on the destination router. Without the second command, no routes appear in the IP routing table because CLNS does not receive the frames to populate the IP routing table. Remember that these are IP routes carried in the IS-IS routing protocol. IS-IS updates the IP routing table.

Example 12-16 *Configuration for Router A: Integrated IS-IS Broadcast Across an NBMA Frame Relay Cloud*

```
Router A
Current configuration:
!
version 12.1
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
```

continues

Example 12-16 *Configuration for Router A: Integrated IS-IS Broadcast Across an NBMA Frame Relay Cloud (Continued)*

```

!
hostname A
!
!
network-clock base-rate 56k
ip subnet-zero
ip tcp synwait-time 5
no ip domain-lookup
!
!
cns event-service server
!
!
controller T1 0
!
!
Interface Ethernet0
 ip address 10.1.128.1 255.255.255.0
 ip router isis
 isis circuit-type level-1
!
Interface Serial0
 ip address 10.100.100.34 255.255.255.0
 ip router isis
 encapsulation frame-relay
 clockrate 56000
 frame-relay map clns 629 broadcast
 frame-relay map clns 931 broadcast
 frame-relay map ip 10.100.100.33 931 broadcast
 frame-relay map ip 10.100.100.35 629 broadcast
 isis circuit-type level-2-only
!
!
router isis
 net 49.0001.0000.0000.000a.00
!
ip classless

```

2. The WAN is a Frame Relay cloud, and Router G has a point-to-point link with Router C. Issue the commands for Router C that configure the link for Integrated IS-IS as a point-to-point link.

The highlighted portion of the configuration file in Example 12-17 shows the configuration of IS-IS across the Frame Relay cloud, using the point-to-point technology and point-to-point Hellos. Because the link is point-to-point, there is no need to configure **frame-relay map** commands (there is no choice of destination to define). The point-to-point link is just a pipe that

goes to one destination. As shown in the configuration, it is only necessary to configure the interface as point-to-point, start Frame Relay, and define the DLCI. In addition to configuring Frame Relay, you must start the IS-IS process for the interface.

Example 12-17 *Configuration for Router C: Integrated IS-IS Point-to-Point Across an NBMA Frame Relay Cloud*

```
Router C

Current configuration:
!
version 12.1
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
!
hostname C
!
!
network-clock base-rate 56k
ip subnet-zero
ip tcp synwait-time 5
no ip domain-lookup
!
cns event-service server
!
!
controller T1 0
!
Interface Ethernet0
  no ip address
  ip router isis
  Shutdown
!
Interface Serial0
  no ip address
  encapsulation frame-relay
  no fair-queue
  clockrate 56000
!
Interface Serial0.1 point-to-point
  ip address 10.100.100.66 255.255.255.0
  ip router isis
  frame-relay interface-dlci 333
!
Interface Serial1
```

continues

Example 12-17 *Configuration for Router C: Integrated IS-IS Point-to-Point Across an NBMA Frame Relay Cloud (Continued)*

```

ip address 10.100.100.35 255.255.255.0
ip router isis
encapsulation frame-relay
clockrate 56000
frame-relay map clns 629 broadcast
frame-relay map clns 631 broadcast
frame-relay map ip 10.100.100.33 631 broadcast
frame-relay map ip 10.100.100.34 629 broadcast
isis circuit-type level-2-only
!
router isis
net 49.0003.0000.0000.000c.00
!
ip classless

```

3. To reduce bandwidth consumption and to hide some network detail, summarization has been suggested as a solution over the WAN links. Issue the commands for Router A that will summarize the networks behind this router with a prefix of /16 across the WAN.

The highlighted portion of the configuration file in Example 12-18 shows the **summary-address** command that is used to hide the routes within area 0001 from the other areas. This configuration is possible on Router A because it sits on the boundary between areas.

Summarizing routes reduces the network resources required by the network.

Example 12-18 *Configuration for Router A: Summarizing Networks from Router A in Area 0001 to All Other Areas*

```

Router A
Current configuration:
!
version 12.1
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
!
hostname A
!
!
network-clock base-rate 56k
ip subnet-zero
ip tcp synwait-time 5
no ip domain-lookup
!
!
cns event-service server

```

Example 12-18 *Configuration for Router A: Summarizing Networks from Router A in Area 0001 to All Other Areas (Continued)*

```

!
!
controller T1 0
!
!
Interface Ethernet0
 ip address 10.1.128.1 255.255.255.0
 ip router isis
 isis circuit-type level-1
!
Interface Serial0
 ip address 10.100.100.34 255.255.255.0
 ip router isis
 encapsulation frame-relay
 clockrate 56000
 frame-relay map clns 629 broadcast
 frame-relay map clns 931 broadcast
 frame-relay map ip 10.100.100.33 931 broadcast
 frame-relay map ip 10.100.100.35 629 broadcast
 isis circuit-type level-2-only
!
!
router isis
 summary-address 10.1.0.0 255.255.0.0
 net 49.0001.0000.0000.000a.00
!
ip classless

```

Scenario 12-2 Answers

1. *Identify the DIS on the Ethernet segment. How was this information apparent?*

On examining the **show interface output** screen in Example 12-19, you can see the circuit ID is 0000.0000.000B.01. This is the system ID of the pseudonode, as is apparent because the last octet has a nonzero value. The system ID of the pseudonode is the system ID of the DIS plus the nonzero octet. Therefore, the DIS is 49.0002.0000.0000.000B.00.

Example 12-19 *Identifying the DIS in Scenario 12-2*

```

Router B
B# show clns interface
Ethernet0 is up, line protocol is up
  Checksums enabled, MTU 1497, Encapsulation SAP
  ERPDUs enabled, min. interval 10 msec.
  RDPDUs enabled, min. interval 100 msec., Addr Mask enabled

```

continues

Example 12-19 *Identifying the DIS in Scenario 12-2 (Continued)*

Congestion Experienced bit set at 4 packets
CLNS fast switching enabled
CLNS SSE switching disabled
DEC compatibility mode OFF for this interface
Next ESH/ISH in 15 seconds
Routing Protocol: IS-IS
Circuit Type: level-1-2
Interface number 0x0, local circuit ID 0x1
Level-1 Metric: 10, Priority: 64, Circuit ID: 0000.0000.000B.01
Number of active level-1 adjacencies: 1
Level-2 Metric: 10, Priority: 64, Circuit ID: 0000.0000.000B.01
Number of active level-2 adjacencies: 1
Next IS-IS LAN Level-1 Hello in 678 milliseconds
Next IS-IS LAN Level-2 Hello in 1 seconds

2. *If Router A died, what would the effect be on the network?*

If Router A died, Routers B and C would not be able to communicate with Router A or with Router D. However, Routers B and C would be able to communicate with each other. The network behind Router A would function, but it would be isolated from the others.

The neighbor tables would fail to hear the Hellos from Router A, and Routers B and C would time out all routes that they had heard from this router. Routers behind Router A would time out Router A from the neighbor table. All the former neighbors of Router A would send LSPs. The LSPs from Router A would be purged from all the databases, new LSPs would flood the network, and the SPF algorithm would be run. Router A and the network behind it would be annexed from the larger network of Routers B and C.

3. *Is summarization possible only on the routers entering the WAN cloud, or is it possible on the networks not shown in the figure, but on the other side of the routers? Give reasons for your answers.*

Summarization is only possible on the Level 1-2 routers, acting in a similar way to an OSPF ABR, and then only if the addressing scheme allows for it to be implemented. The Level 1-2 router would summarize the routes and inject them into the Level 2 network.

Part IV: EIGRP

Chapter 13 Using EIGRP in Enterprise Networks

Chapter 14 Configuring EIGRP in Enterprise Networks

Part IV covers the following Cisco BSCI exam topics:

- Describe the features and operation of EIGRP
- Given a set of network requirements, identify the steps to configure an EIGRP environment and verify proper operation (within described guidelines) of your routers
- Identify the steps to verify EIGRP operation
- Interpret the output of various **show** and **debug** commands to determine the cause of route selection errors and configuration problems



This chapter covers the following topics, which you need to understand to pass the CCNP/CCDP/CCIP BSCI exam:

- Overview of EIGRP in an enterprise network
- The operation of EIGRP
- EIGRP network design

Using EIGRP in Enterprise Networks

This chapter covers the Enhanced Interior Gateway Routing Protocol (EIGRP) in detail. Although EIGRP has the capability of supporting IP, AppleTalk, and IPX, the BSCI exam will deal with only the mechanics of the IP routing protocol. This chapter expands on your understanding of routing within large enterprise networks, which was covered in Chapter 8, “Using OSPF Across Multiple Areas.”

This chapter deals with how EIGRP works theoretically. The operation of EIGRP, some of the options available, and design considerations are explained in this chapter, particularly in reference to scaling EIGRP and its use over a nonbroadcast multiaccess (NBMA) WAN environment.

The topics in this chapter directly reflect questions on the BSCI exam. EIGRP is designed for use in large networks. As a proprietary routing protocol for Cisco, it is therefore an obligatory subject in a Cisco exam on IP routing protocols.

“Do I Know This Already?” Quiz

The purpose of the “Do I Know This Already?” quiz is to help you to decide what parts of this chapter to use. If you already intend to read the entire chapter, you do not necessarily need to answer these questions now.

The 12-question quiz, derived from the major sections in the “Foundation Topics” portion of the chapter, helps you to determine how to spend your limited study time.

Table 13-1 outlines the major topics discussed in this chapter and the “Do I Know This Already?” quiz questions that correspond to those topics.

Table 13-1 “Do I Know This Already?” Foundation Topics Section-to-Question Mapping

Foundation Topics Section	Questions Covered in This Section
Overview of EIGRP in an Enterprise Network	1–3
The Operation of EIGRP	4–8
EIGRP Network Design	9–12

NOTE The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark this question wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

1. Which of the following describes how a reply to a query is sent?
 - a. Multicast
 - b. Best effort unicast
 - c. A reliable unicast
 - d. A reliable multicast
2. Which of the following are advantages to using EIGRP as a routing protocol?
 - a. Rapid convergence
 - b. Reduced bandwidth consumption
 - c. Support for other Layer 3 protocols
 - d. Increased vendor support
3. Cisco identifies four main components to EIGRP. Select the two correct components listed below.
 - a. Neighbor discovery
 - b. SPF algorithm
 - c. Areas
 - d. RTP
4. For EIGRP to form a neighbor relationship, which of the following must be true?
 - a. Authentication must be enabled.
 - b. The k-values of the metric must be the same on both routers.
 - c. The autonomous system number must be the same on both routers.
 - d. The holddown timer must be the same.
5. The neighbor table contains which of the following information?
 - a. Records sequence numbers
 - b. Lists the metric of neighbors

- c. Calculates the feasible distance between neighbors
 - d. Displays the status of the links between neighbors
- 6. What does the initialism SRTT stand for?
 - a. Shortest Remote-Trip Time
 - b. Smooth Round-Trip Time
 - c. Shortest Reliable-Trip Time
 - d. Single Remote Timer Test
- 7. Which of the following would trigger the topology table to recalculate?
 - a. LSP received
 - b. SRT packet received
 - c. A new router coming online
 - d. Link lose detected
- 8. What do the letters SIA stand for?
 - a. Stuck in Active
 - b. Shortest IP Address
 - c. Stuck in Area
 - d. Simple IP Address
- 9. EIGRP is designed for use in which kind of network?
 - a. Multivendor networks
 - b. ISPs
 - c. Small networks, particularly stub networks across an NBMA cloud
 - d. Large networks
- 10. Why might queries stop at the point where summarization is configured? Select the best answer.
 - a. Summarization stops the query process.
 - b. Summarization might hide the subnet that is being queried.
 - c. Queries do not travel across network boundaries.
 - d. Queries do not travel across areas.

11. What can be done to prevent feedback loops during redistribution between EIGRP autonomous systems?
 - a. Disabling the query process
 - b. Manually configuring the query holdtime
 - c. Applying route filters
 - d. Implementing summarization

12. Which of the following reasons can lead to a poorly scaled network?
 - a. The number of EIGRP processes
 - b. The number of alternative paths to remote networks
 - c. The number of times that RTP needs to be run
 - d. The SIA algorithm configuration

The answers to this quiz are found in Appendix A, “Answers to Chapter ‘Do I Know This Already?’ Quizzes and Q&A Sections.” The suggested choices for your next step are as follows:

- **6 or less overall score** —Read the entire chapter. This includes the “Foundation Topics” and “Foundation Summary” sections, and the “Q&A” section at the end of the chapter.
- **7–9 overall score** —Begin with the “Foundation Summary” section, and then go to the “Q&A” section at the end of the chapter. If you have trouble with these exercises, read the appropriate sections in “Foundation Topics.”
- **10 or more overall score** —If you want more review on these topics, skip to the “Foundation Summary” section, and then go to the “Q&A” section at the end of the chapter. Otherwise, move to the next chapter.

Foundation Topics

Overview of EIGRP in an Enterprise Network

EIGRP is an enhanced version of IGRP, hence the name. It uses the same distance vector technology as IGRP. The changes were effected in the convergence properties and the operating efficiency of the protocol. EIGRP has some characteristics similar to those of a link-state routing protocol. Therefore, it is sometimes referred to as a *hybrid routing protocol*, although Cisco calls it an *advanced distance vector protocol*. EIGRP is an efficient, although proprietary, solution to networking large environments because it scales well. Its ability to scale is, like OSPF, dependent on the design of the network.

EIGRP Terminology

To understand how EIGRP works, you must be familiar with the terminology. Table 13-2 defines the main components and concepts.

Table 13-2 *Terminology for EIGRP for IP*

Term	Definition
Neighbor	A router running EIGRP that is directly connected.
Neighbor table	A list of every neighbor, including the IP address, the outgoing interface, the holdtime, smooth round-trip time (SRTT), and uptime, or how long since the neighbor was added to the table. This table is built from information on Hellos received from adjacent routers (neighbors).
Route table	The routing table, or list of available networks and the best paths. A path is moved from the topology table to the routing table when a feasible successor is identified.
Topology table	A table that contains all the paths advertised by neighbors to all the known networks. This is a list of all the successors, feasible successors, the feasible distance, the advertised distance, and the outgoing interface. DUAL acts on the topology table to determine successors and feasible successors by which to build a routing table.
Hello	Messages used to find and maintain neighbors in the topology table. They are sent periodically and are sent unreliably.

(continues)

Table 13-2 *Terminology for EIGRP for IP (Continued)*

Term	Definition
Update	An EIGRP packet containing change information about the network. It is sent reliably. It is sent only when there is a change in the network to affected routers: • When a neighbor first comes up • When a neighbor transitions from active to passive for a destination • When there is a change in calculated metric for a destination
Query	Sent from the router when it loses a path to a network. If there is no alternate route (feasible successor), it will send out queries to neighbors inquiring whether they have a feasible successor. This makes the route state change to active. The queries are sent reliably.
Reply	A response to the query. If a router has no information to send in a reply, it will send queries to all its neighbors. A unicast is sent reliably.
ACK	A Hello packet with no data that is an acknowledgment of packets sent reliably.
Holdtime	Value set in the Hello packet. It determines how long the router waits for Hellos from a neighbor before declaring it unavailable. This information is held in the neighbor table.
Smooth Round-Trip Time (SRTT)	The time that the router waits after sending a packet reliably to hear an acknowledgment. This is held in the neighbor table and is used to calculate the RTO.
Retransmission Timeout (RTO)	Timer calculated in reference to the SRTT. RTO determines how long the router waits for an ACK before retransmitting the packet.
Reliable Transport Protocol (RTP)	Mechanism used to determine requirements that the packets be delivered in sequence and guaranteed.
Diffusing Update Algorithm (DUAL)	An algorithm performed on the topology table to converge the network. It is based on a router detecting a network change within a finite time, with the change being sent reliably and in sequence. As the algorithm is calculated simultaneously, in order, and within a finite time frame on all affected routers, it ensures a loop-free network.
Advertised distance (AD)	The cost of the path to the remote network from the neighbor (the metric from the next-hop router).
Feasible distance (FD)	The lowest-cost distance (metric) to a remote network.

Table 13-2 *Terminology for EIGRP for IP (Continued)*

Term	Definition
Feasible condition (FC)	When a neighbor reports a path cost (AD) that is lower than the router's FD to a network. The neighbor's (next-hop router's) path has a lower metric than the router's path.
Feasible successor (FS)	The neighbor reporting the AD that is lower than the router's FD becomes the feasible successor. The next-hop router that meets the FC.
Successor	The next-hop router that passes the FC. It is chosen from the FSs as having the lowest metric to the remote network.
Stuck in Active (SIA)	State reached when a router has sent out network packets and is waiting for ACKs from all its neighbors. The route is active until all the ACKs have been received. If they do not appear after a certain time, the router is SIA for the route.
Query scoping	Network design to limit the scope of the query range, that is, how far the query is allowed to propagate in search of a feasible successor. This is necessary to prevent SIA, which can cause multiple problems for the network.
Active	Route state when there is a network change, but after examining the topology table, no FS is found. The route is set to active mode, and the router queries its neighbors for alternative routes.
Passive	An operational route is passive. If the path is lost, the router examines the topology table to find an FS. If there is an FS, it is placed in the routing table; otherwise, the router queries its neighbors, sending the route into active mode.

Understanding EIGRP Features and Advantages

The goal of EIGRP is to solve the scaling limitations that IGRP faces, using the distance vector technology from which it grew. EIGRP increases the potential growth of a network by reducing the convergence time. This is achieved by the following features:

- DUAL
- Loop-free networks
- Incremental updates
- Multicast addressing for updates
- Advanced distance vector protocol

- Loop-free routing tables
- Support for different topologies
- Rapid convergence
- Reduced bandwidth use
- Protocol independence at Layer 3
- Compatibility with IGRP
- Easy configuration
- Use of a composite metric
- Unequal-cost load balancing

A full understanding of the concepts and operation of EIGRP will aid you in the design, implementation, and maintenance of EIGRP networks and will definitely help you pass an exam on the subject. The following sections describe each of these features in more detail.

DUAL

DUAL is one of the main features of EIGRP. It diffuses the routing computation over multiple routers. A more detailed discussion of DUAL is provided in the section “Components of EIGRP.”

Loop-Free Networks

The DUAL algorithm is used to ensure a loop-free network. The calculation of an FS means that the backup route is downstream from the router. The FS is chosen only because it has a lower metric to the destination route than that reported by the router. This prevents any routes that lead back to the router from being chosen, thus eliminating loops.

Incremental Updates

EIGRP sends nonperiodic, partially bounded updates. That is, they are sent whenever there is a change to be reported and not at regular intervals. When a network change is made, the updates include only those changes that are needed to update the affected routers.

Multicast Addressing for Updates

EIGRP uses an RTP that guarantees delivery. This is essential when the routing updates are not sent periodically; otherwise, if the receiving router is not expecting an update, it cannot realize that an update was missed, indicating a network problem. Updates are therefore sent using a reliable multicast. The address is the reserved class D address, 224.0.0.10. When the neighbor receives a multicast, it acknowledges receipt of the packet with an unreliable unicast.

Advanced Distance Vector Protocol

EIGRP has solved many of the problems inherent in distance vector protocols, which prevent them from effectively supporting large networks. The features of distance vector technology that prohibit network growth include the use of broadcasts and hop count, neither of which is used by EIGRP. Other characteristics that advance EIGRP beyond an ordinary distance vector protocol come from its status as a classless routing protocol. Without the use of areas, EIGRP allows summarization anywhere in the network. Summarization reduces the need for network resources. Because the classless protocols send the subnet mask with the update, this also means that classless protocols support discontinuous networks and, of course, variable-length subnet mask (VLSM).

Loop-Free Routing Tables

The criteria for selecting the primary and backup routes in the topology table and the routing table ensure that the routes offered are loop-free. The primary route that is placed in the routing table is chosen for the lowest metric, which means it cannot be looped. The backup route (feasible successor) is dependent on the downstream router (next hop advertising the alternative route) advertising a lower cost for the route than the one stored in the routing table. This ensures that the backup route does not loop back through the router.

Support for Different Topologies

EIGRP, as a new protocol, has been able to anticipate recent topologies, such as NBMA clouds. There is no complex configuration required for these topologies, though additional configuration is available for tuning the update operation of EIGRP.

Rapid Convergence

The use of the DUAL algorithm stores not only the best path to the destination, but also the close contenders. If a network fails, the router can immediately switch to the alternate route. If there are no alternative routes, then the router will query neighbors to see whether they have a path to the destination.

Reduced Bandwidth Use

Using multicast and unicast addressing to send and acknowledge updates restricts the potential use of both bandwidth and the other system's CPU to the essential requirements. EIGRP also uses only incremental updates, as opposed to periodic updates.

Protocol Independence at Layer 3

EIGRP functions as the routing protocol for IP, AppleTalk, and IPX. A different routing table is maintained for each Layer 3 protocol. EIGRP will automatically redistribute IPX RIP, AppleTalk RTMP, and IP IGRP within the same autonomous system.

Compatibility with IGRP

Because it grew out of IGRP, EIGRP is backward-compatible with IGRP. This allows for seamless transitions to EIGRP and support for older, smaller networks that have neither the need nor the capability to upgrade. EIGRP automatically redistributes IP routes learned into the IGRP process as long as the autonomous system number used to configure the processes is the same.

Easy Configuration

Because EIGRP was designed for the hardware on which it runs, the protocol not only is tuned for efficiency, but also is simple and straightforward to configure. Another benefit is that EIGRP has fewer design constraints than OSPF; for example, EIGRP supports point-to-point, in addition to NBMA point-to-point and multipoint. EIGRP requires no additional configuration other than tuning the bandwidth utilization, if desired.

Use of a Composite Metric

EIGRP uses the same metric as IGRP (bandwidth and delay as the default), though EIGRP has expanded the metric to 32-bit, allowing for greater scaling and granularity. An intelligent metric will select the shortest path.

Unequal-Cost Load Balancing

Unequal-cost load balancing allows all links to a destination to be used to carry data without saturating the slower links.

NOTE A broadcast domain identifies devices that are within the same Layer 2 domain. Although they might not be directly connected to the same physical cable, if they are in a switched environment, from a logical Layer 2 or Layer 3 perspective, they are on the same link. If a broadcast is sent out, all the devices within the broadcast domain will hear the message and will expend resources determining whether it is addressed to them. A Layer 3 device is a broadcast firewall in that a router does not forward broadcasts.

Components of EIGRP

Cisco identifies four main components of EIGRP:

- Protocol-dependent modules
- RTP
- Neighbor discovery and recovery

The following sections describe each of these components in more detail.

Protocol-Dependent Modules

EIGRP functions as the routing protocol for IP, AppleTalk, and IPX. A different routing table is maintained for each Layer 3 protocol.

RTP

EIGRP uses both multicast and unicast addressing. Some of these packets are sent reliably using RTP, meaning that they require an acknowledgment. These packets are also sent with sequence numbers to make the transmission of data reliable. Hellos and ACKs do not require acknowledgment, but the incremental updates cannot be anticipated. Therefore, the update, query, and reply packets must be acknowledged by the receiving neighbor or the packet is retransmitted up to 16 times.

Neighbor Discovery and Recovery

EIGRP operates between neighbors that share routing tables and information about the state of their connections. In this way, the routing protocol localizes as much information as possible, reducing the bandwidth and CPU requirements of the network and speeding up convergence.

DUAL

EIGRP uses DUAL to maintain the network databases. It selects the shortest path to a destination and then maintains a backup path if available, making network convergence almost instantaneous. Using the terminology for EIGRP, the term *successor* refers to the path to a destination. The successor is chosen using DUAL from all of the known paths or feasible successors to the end destination. If the selected path dies for any reason, an alternative route is chosen from the feasible successors. If a feasible successor is found, the route stays in passive mode, and there is no disruption to the network. However, if an alternative path is not immediately available, the router queries its neighbors, placing the route in active mode.

The Operation of EIGRP

One of the main strengths of EIGRP is that it limits the scope of network computation, keeping all knowledge of network change as local as possible. EIGRP is a protocol that works on a “need to know” basis.

Even if the computation of the network is local, the router must know about the entire network. The explanation of the routing protocol is given through the viewpoint of one router. Once you understand the network communication between the routers running EIGRP, the operation of EIGRP becomes clear; the concepts and terms are placed in context. This facilitates your understanding of the subject; rote memorization is no longer necessary.

There are three main tables in EIGRP:

- The neighbor table
- The topology table
- The routing table

Creating the Neighbor Table

The neighbor table is maintained by means of the Hello protocol, a small packet sent out to dynamically learn of other routing devices that are in the same broadcast domain. It informs neighbors that connections are live and active and keeps track of all the packets sent between the neighbors. This table is therefore maintained not only by the Hello protocol, but also by actively monitoring all traffic between the directly connected systems.

Each Layer 3 protocol supported by EIGRP (IP, IPX, and AppleTalk) has its own neighbor table, which makes sense because the neighbor, topology, and routing tables differ greatly. Although all the information is held in one table, the different EIGRP processes all have to access the same table, which complicates and slows down the lookup.

The Contents of the Neighbor Table

The neighbor table includes the following information:

- The address of the neighbor.
- The interface through which the neighbor's Hello was heard.
- The holdtime, or how long the neighbor table waits without hearing a Hello from a neighbor, before declaring the neighbor unavailable and purging the database. This is three times the value of the Hello timer by default.
- The uptime, or how long since the router first heard from the neighbor.
- The sequence number. The neighbor table tracks all the packets sent between the neighbors. It tracks both the last sequence number sent to the neighbor and the last sequence number received from the neighbor. Although the Hello protocol is a connectionless protocol, other protocols used by EIGRP are connection-oriented. The sequence number is in reference to these protocols.
- SRTT. This calculates the RTO. This is the time in milliseconds that it takes a packet to be sent to a neighbor and a reply to be received.
- RTO. This states how long the router will wait on a connection-oriented protocol without an acknowledgment before retransmitting the packet. If the original packet that was unacknowledged was multicast, the retransmitted packets will be unicast.

- The number of packets in a queue. This is a means by which administrators can monitor congestion on the network.

Becoming a Neighbor

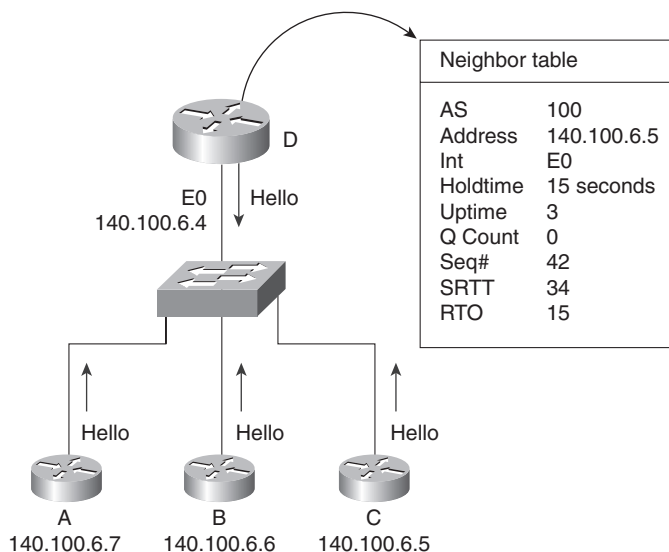
The EIGRP Hello protocol uses a multicast address of 224.0.0.10, and all routers periodically send Hellos. On hearing Hellos, the router creates a table of its neighbors. The continued receipt of these packets maintains the neighbor table. If a Hello from a known neighbor is not heard within a predetermined amount of time, as stated in the holdtime, the router will decide that the neighbor is no longer operational and will take the appropriate action. The holdtime is set at the default of three times the Hello timer. Therefore, if the router misses three Hellos, the neighbor is declared dead. The Hello timer on a LAN is set to 5 seconds; the holdtime, therefore, is 15 seconds. On a WAN link, the Hello timer is 60 seconds, and the holdtime correspondingly is 180 seconds.

To become a neighbor, the following conditions must be met:

- The router must hear a Hello packet or an ACK from a neighbor.
- The autonomous system number in the packet header must be the same as that of the receiving router.
- The neighbor's metric settings must be the same as that of the receiving router.

Figure 13-1 demonstrates building the neighbor table.

Figure 13-1 *Building the Neighbor Table*



Creating the Topology Table

After the router knows who its neighbors are, it is able to create a database of feasible successors. The neighbors and available feasible successors are held in the topology table. Contrary to popular belief, the topology table has a record of all known network routes within the organization, not simply the feasible successors and successors. The other routes are referred to as *possibilities*. The topology table in EIGRP manages the selection of routes to be added to the routing table.

The topology table includes the following information:

- Whether the route is passive or active.
- That an update has been sent to the neighbors.
- That a query packet has been sent to the neighbors. If this field is positive, at least one route will be marked as active.
- If a query packet has been sent, another field will track whether any replies have been received from the neighbors.
- That a reply packet has been sent in response to a query packet received from a neighbor.
- The remote networks.
- The prefix or mask for the remote network.
- The metric for the remote network, the FD.
- The metric for the remote network advertised by the next logical hop, the AD.
- The next hop.
- The outgoing interface to be used to reach the next logical hop.
- The successors, the path to the remote network stated in hops.

The table is built from the update packets that are exchanged by the neighbors and by replies to queries sent by the router. Replies are sent in response to queries, inquiring about suspect routes.

The queries and responses used by EIGRP for DUAL are sent reliably as multicasts using RTP, which was created by and is proprietary to Cisco. If a router does not hear an acknowledgment within the allotted time, it retransmits the packet as a unicast. If there is no response after 16 attempts, the router marks the neighbor as dead. Each time the router sends a packet, RTP increments the sequence number by one. The router must hear an acknowledgment from every router before it can send the next packet. The capability to send unicast retransmissions decreases the time that it takes to build the tables.

When the router has an understanding of the network, it runs DUAL to determine the best path to the remote network. The result is entered into the routing table.

Maintaining the Topology Table

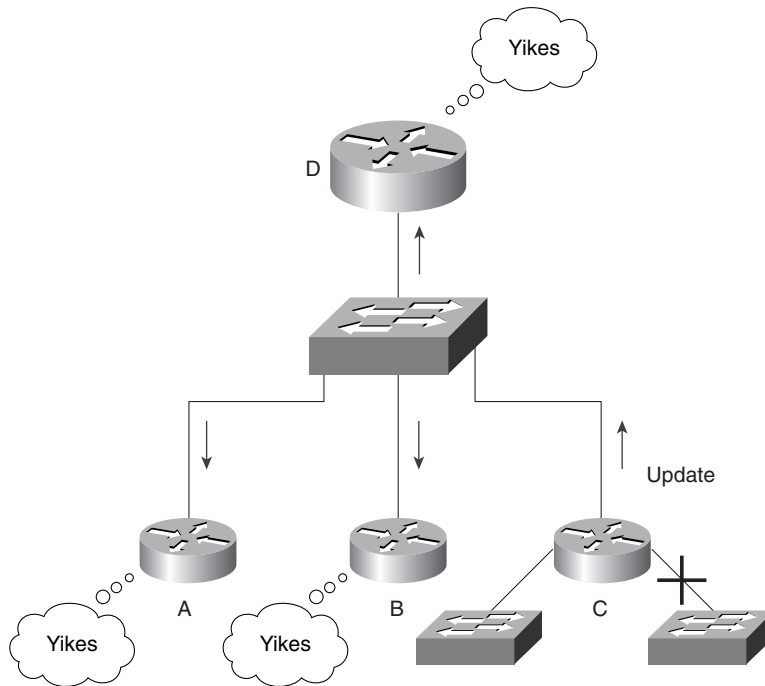
The following three reasons might cause a topology table to be recalculated:

- The router hears a change when a new network is available because of one of the following reasons:
 - The topology table receives an update stating that there is a new remote network.
 - The interface for a directly connected EIGRP network comes online.
- The router changes the successor in the topology table and routing table in these circumstances:
 - The topology table receives a reply or a query from a neighbor.
 - There is local configuration of a directly connected interface to change the cost of the link.
- The router hears a change from a neighbor when a network has become unavailable because of one of the following reasons:
 - The topology table receives a query, reply, or update stating that the remote network is down.
 - The neighbor table does not receive a Hello within the holdtime.
 - The network is directly connected, and the router senses a loss of carrier.

Figure 13-2 illustrates the traffic flow seen when a router loses a direct connection.

Just as the neighbor table tracks the receipt of the EIGRP packets, the topology table records the packets that have been sent by the router to the neighbors. It also identifies the status of the networks in the table. A healthy network is marked as *passive*; it is labeled as *active* if the router is attempting to find an alternative path to the remote network that is believed to be down.

Because the routing table is built from the topology table, the topology table must have the information required by the routing table. This includes the next hop, or the address of the neighbor that sent the update with that network. The routing table also calculates the metric to the remote network.

Figure 13-2 *Maintaining the Topology Table—the Traffic Flow*

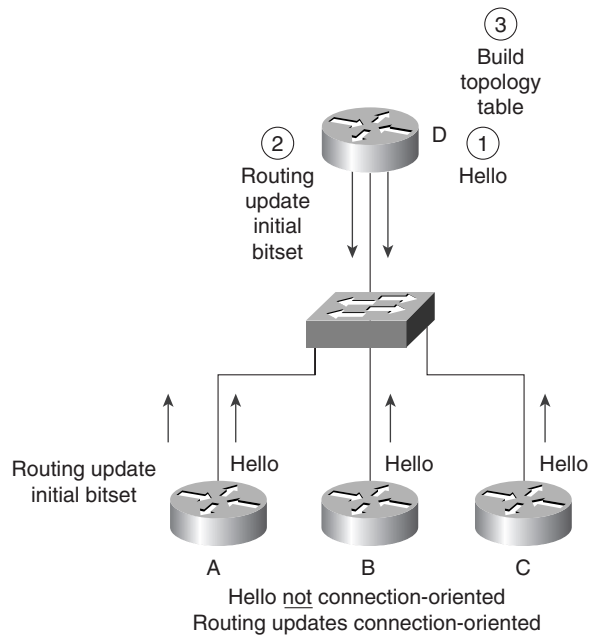
Adding a Network to the Topology Table

Figure 13-3 shows the topology for this section.

Imagine an access layer router (Router A) that connects to a new network. The administrator has connected and configured another Ethernet interface to service a department that has moved into the building. The following list describes how the new network is propagated to all the routers in the EIGRP autonomous system:

1. As soon as Router A becomes aware of the new network, it starts to send Hello packets out of the new interface. No one answers because this is an access router giving connectivity to the workstations and other end devices.

There are no new entries in the neighbor table because no neighbors have responded to the Hello protocol. There is a new entry in the topology table, however, because this is a new network.

Figure 13-3 *EIGRP—Updating the Topology Table with a New Router*

2. EIGRP, sensing a change, is obliged to send an update to all its neighbors, informing them of the new network. The updates set a bit in the flag field of the update header. The initial bit shows that the update includes route entries that are the first in the setup of a new neighbor relationship. These updates are tracked in the topology table and the neighbor table because the updates are connection-oriented and the acknowledgments from the neighbors must be received within a set timeframe.

Router A, having added the network to its topology table, adds the network to its routing table. The network will be marked as passive because it is operational.

Router A's work is done. Router D's work has just begun. Acting at the distribution layer, Router D is the router that connects Routers A, B, and C to the rest of the building. Its neighbors are routers on each floor and the routers in the other buildings.

3. On hearing the update from Router A, Router D updates the sequence number in the neighbor table and adds the network to the topology table. It calculates the FD and the successor to place in the routing table. It is then in a position to send an update to all of its neighbors, except for Routers A, B, and C. It is obeying the split horizon rule here. Routers B and C are updated in the same manner and at the same time as Router D.

The next section describes the process for removing a router or path from the topology table.

Removing a Path or Router from the Topology Table

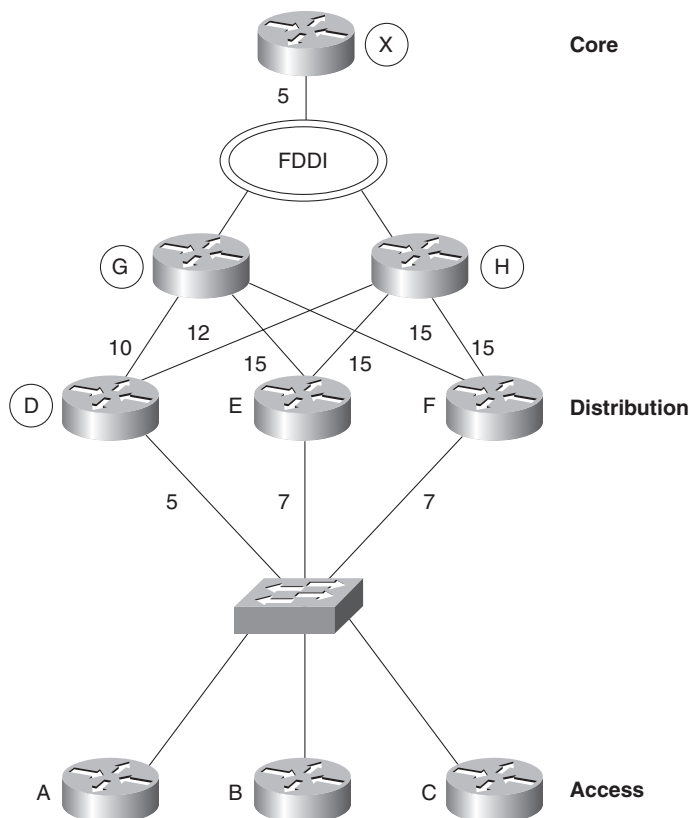
The process of removing a path or router from the topology table is far more complex and gets to the crux of EIGRP. The following process uses Figure 13-3 and focuses on Router D:

1. If a network connected to Router A is disconnected, Router A updates its topology and routing table and sends an update to its neighbors.
2. When Router D receives the update, it updates the neighbor table and the topology table.
3. As a router, D is programmed to find an alternative route to the remote network. It examines the topology table for alternatives. Because there is only one path to the remote network, no alternatives are found.
4. The router then sends out a query to its neighbors requesting that they look in their tables for paths to the remote network. The route is marked active in the topology table at this time.
5. The query is tracked, and when all the replies are in, the neighbor and topology tables are updated.
6. DUAL, which starts to compute as soon as a network change is registered, runs to determine the best path, which is placed in the routing table.
7. Because no alternative route is available, the neighbors reply to the query stating that they have no path.
8. Before they respond, they query their own neighbors; in this way, the search for an alternative path extends or diffuses throughout the organization.
9. When no router can supply a path to the network, all the routers remove the network from their routing and topology tables.

The next section describes when a neighbor does have an alternative route.

Finding an Alternative Path to a Remote Network

When the path to a network is lost, EIGRP goes to a lot of trouble to find an alternative path. This process is one of the major benefits of EIGRP. The method that EIGRP uses to find alternative paths is very reliable and very fast. Figure 13-4 illustrates the steps in this section. In the figure, the routers that are participating in the selection process are circled.

Figure 13-4 Campus Topology Map Showing Alternative Path Selection

NOTE The metric shown in Figure 13-4 has been simplified for the purposes of this example.

The following list describes the process after Router G goes offline:

1. Router D marks the routes that were reached by sending the traffic to Router G.
2. Router D looks in the topology table, which has every network and path of the network, to determine whether there is an alternative route. It is looking for an FS.
3. An FS is determined by a clearly defined equation. The topology table has an AD and an FD listed for every route or successor. This comprises the metric by which the route was selected.
4. Router D adds the alternative route to Router X via Router H, found in the topology table, without moving into active mode because the AD is still less than the original FD. The AD is 5; the original FD was 15. It needs to send updates to its neighbors because the distance has changed.

5. If the router did not have an FS, it would have placed the route into an active state while it actively queried other routers for an alternative path.
6. After interrogating the topology table, if a feasible route is found, the neighbor replies with the alternative path. This alternative path is then added to the topology table.
7. Next, in the last steps of DUAL, the routing table is updated.
8. The network is placed back into a passive state as the router returns to the normal forwarding and maintenance of EIGRP tables until the next change in the network.
9. If a neighbor that has been queried has no alternative path or FS, it places the network into active mode and queries its neighbors.
10. If no answer is heard, the messages are propagated until they hit a network or autonomous system boundary.

When the router sends a query packet, it is recorded in the topology table. This is to ensure a timely reply. If the router does not hear a reply, the neighbor is removed from the neighbor table; all the networks held in the topology table for that neighbor are seen as suspect, and the networks are queried. Occasionally, because of slow links and burdened routers in a large network, problems can occur. In particular, a router might not receive a reply from all the queries that it sent out. This leads to the route being declared SIA; the neighbor that failed to reply is removed from the neighbor table, and DUAL assumes a reply was received, giving an infinite metric.

NOTE SIA can take minutes to resolve. First, it is important to discover if neighbors are consistently failing to answer the query and why. This failure is due to a resource problem either on the router or on the link to the router. It is always better to redesign the network. Summarization and distribution (route filters) both reduce the scope of the query range. This subject is well documented in the Cisco White Paper “EIGRP—Enhanced Interior Gateway Routing Protocol” available at the Cisco Web site, Cisco.com.

Creating the Routing Table

The routing table is built from the topology table after DUAL has been run. The topology table is the foundation of EIGRP: This is where all the routes are stored, even after DUAL has been run. It is in the routing table that the best paths are stored and accessed by the routing process.

Once the tables have been built, the router can make routing decisions.

EIGRP Metrics

The metrics used in EIGRP are very similar to those of IGRP. The main difference is that the result of the calculation is held in a 32-bit field. This means that the decision can be much finer or more detailed. The DUAL algorithm uses this metric to select the best path or paths to a destination. The

computation is performed on paths held in the topology table to identify the best path to place into the routing table. Up to six paths can be held for one destination, and there can be three different types of paths. These three path types are described in Table 13-3.

Table 13-3 *EIGRP Routing Types*

Route Type	Description
Internal	Internal paths to the autonomous system
Summary	Internal paths that have been summarized
External	External paths to the autonomous system that have been redistributed into this EIGRP autonomous system

The metric is the same composite metric used by IGRP, with the default calculated from bandwidth and delay. Although it is possible to change the metric, this must be done only with great care and consideration to the network design. Any changes made must be effected on every router in the EIGRP autonomous system.

The equation for the default metric used is this:

$$\text{metric} = [(10000000 \mid \text{smallest bandwidth kbps}) + \text{sum of delays}] * 256$$

Table 13-4 explains the metric values.

Table 13-4 *EIGRP Metric Values*

Metric Symbol	Metric Value	Description
K1	Bandwidth	Selects the smallest bandwidth media between the source and destination hosts. The equation used is $[10000000 \mid \text{bandwidth kbps}] * 256$.
K2	Loading	Based on the statistics held at the outgoing interface and recorded in bits per second.
K3	Delay	The delay calculated on the outgoing interface. The value used is the cumulative total of the delay on all the interfaces between the hosts. The delay is measured in units of 10 ms to 168 seconds. A delay of all 1s in the 32-bit field means the network is unreachable.
K4	Reliability	Based on the statistics held on the outgoing interface gained from keepalives and exponentially averaged over five minutes.
K5	MTU	The smallest MTU found on an interface along the route. This value is included although it has not been used as part of the metric calculation.

The default for the K constants are

$$K1 = 1, K2 = 0, K3 = 1, K4 = 0, K5 = 0$$

If $K5 = 0$, the composite metric is calculated using the following formula:

$$\text{metric} = [K1 * \text{bandwidth}(K2 * \text{bandwidth})(256 - \text{load}) + K3 * \text{delay}]$$

If $K5$ is not 0, the added formula is used:

$$\text{metric} = \text{metric} * [K5(\text{reliability} + K4)]$$

Given the overall understanding of how EIGRP works, the next section considers the topology table and its components, which will help explain the details of EIGRP operation.

The Topology Table and the DUAL Finite-State Machine

DUAL is responsible for maintenance of the topology table and the creation of the routing table. The topology table records the metric as received from the advertising router, or the next hop. It then adds the cost of getting to that neighbor, the one that is advertising the route. The cost to the destination network from the advertising router, plus the cost to that router, equals the metric to the destination network from the router.

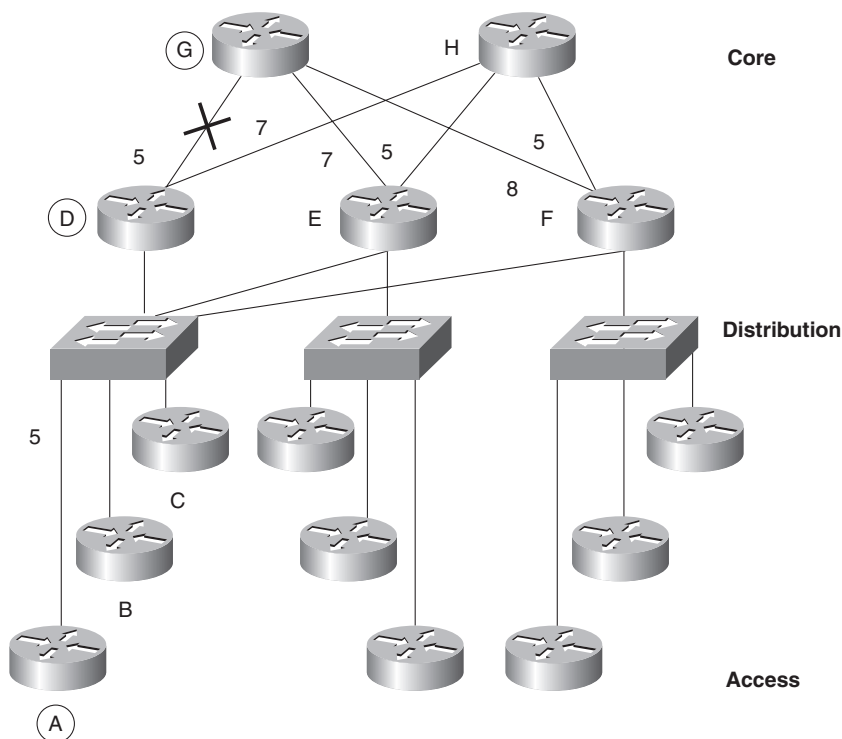
The metric or cost from the neighbor advertising the route is known as the *advertised distance (AD)*. The metric or cost from the router that is determining the metric or the local router is referred to as the *feasible distance (FD)*. If the AD is less than the FD, the next-hop router is downstream and there is no loop. Put simply: The downstream neighbor or next hop must be closer to the destination. This is fundamental to EIGRP.

Figures 13-4 and 13-5, shown earlier in this chapter, illustrate these distances. Note that the metric shown in these figures has been simplified for the purposes of this example.

Updating the Routing Table in Passive Mode with DUAL

DUAL determines whether there is an acceptable route in the topology table to replace the current path in the routing table. In EIGRP terms, this is replacing a successor in the routing table with a feasible successor from the topology table.

Use the network in Figure 13-5 as an example.

Figure 13-5 *The Use of Feasible and Advertised Distance—Passive Mode*

The following list explains the figure with the metrics and actions that EIGRP takes in determining the path:

- The FD from Router A to Router G is 10 (A to D to G).
- The AD from Router A to Router G is 5 (advertised from Neighbor D).
- Because $10 > 5$, then $FD > AD$. This means that the FD is a feasible condition (FC), allowing it to become an FS. If you follow the diagram, it is very straightforward and less algebraic.
- If the link between Router D and Router G were down, Router A would look in its topology table.
- The alternative routes through Router A to D to H to E to G have an AD of 19 ($7 + 5 + 7$).
- Because 19 is greater than the original FD of 10, it does not qualify as an FS.
- The path through Router D to H to F to G has an AD of 20 and cannot be an FS.

- The path through Router A to E to G has an AD of 7, however, which is less than the original 10. Therefore, this is an FS and can be replaced as a route without Router A changing from passive to active mode.
- The original topology table would show that the primary route (successor) is Router D, while the backup route (FS) is Router E. After the link between D and G dies, the routing table would be updated from the topology table while the route remains passive.

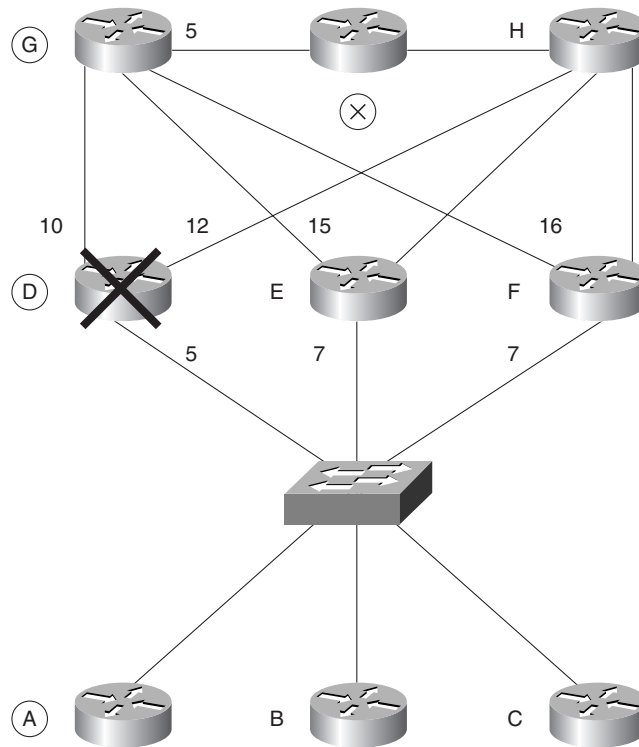
The following section illustrates what happens when the topology table is interrogated and no feasible route is found.

Updating the Routing Table in Active Mode with DUAL

When no alternative route is found in the routing table, the following actions are taken (using the network in Figure 13-6 as an example). The following list describes the figure and explains the actions taken on the information provided:

- The topology table of Router A has a path (successor) of A to D to G to X.
- The FD is 20, and the AD from Router D is 15.
- When Router D dies, Router A must find an alternative path to X.
- Neighbors B, C, E, and F have ADs of 27, 27, 20, and 21, respectively.
- Because all the neighbors have an AD that is the same or greater than the successor FD, none of these are acceptable as FSs.
- Router A must go into active mode and send queries to the neighbors.
- Both Routers E and F reply with an FS because both have an AD from Router G of 5. Remember the equation $FD > AD$; the Routers E and F have an FD of 21, and $21 > 5$.
- Because the FD is acceptable, the topology and routing tables will be updated, DUAL will be calculated, and the network will be returned to passive mode.
- From this information received from Routers E and F, the router selects the path through E as the best route because it has the lower cost.
- The result is placed in the routing table as the valid neighboring router. EIGRP refers to this neighboring router as a *successor*.
- Router F will be stored as an FS in the topology table.

NOTE Figure 13-6 is simplified to explain the concepts. In reality, the split horizon rule dictates that Routers B and C would not readvertise routes it learned through an interface out of that same interface. Because all routes to X are learned through one interface, no routes to X would be readvertised out of this interface.

Figure 13-6 *The Use of Feasible and Advertised Distance—Active Mode*

The details on how EIGRP computes successors are complex, but the concept is simple, as described in the next section.

Choosing a Successor

To determine whether a path to a remote network is feasible, EIGRP considers the feasible condition (FC) of the route. Essentially, each router holds a routing table that is a list of the available networks and the best or most efficient path to each of them. The term used to describe this is the *feasible distance of the successor*, otherwise known as the metric for the route. The router also holds the routing table of its neighbors, referred to as the AD. If the AD is within scope, this route may be identified as an alternative route, or an FS.

A neighbor can become an FS for a route only if its AD is less than the FD. This is DUAL's fundamental key to remaining loop-free; if a route contains a loop, the AD will be greater than the FD and therefore will fail the FC. By holding the routing tables of the neighbors, the amount of network overhead and computation is reduced. When a path to a remote network is lost, the router

might be capable of finding an alternative route with minimal fuss, computation, or network traffic. This gives the much-advertised benefit of very fast convergence.

As you can see in the explanation for finding an FS in the previous section, “Updating the Routing Table in Active Mode with DUAL,” queries can be sent throughout the organization’s network. This is the design key to ensuring that EIGRP scales.

EIGRP Network Design

EIGRP is designed to work in very large networks. However, EIGRP, as with OSPF, is design-sensitive. Scaling a network—or, in other words, improving its capability to grow in size and complexity—is a major concern in today’s organizations. New demands are constantly driving the networks to use applications that require more bandwidth and other resources from the network. For example, simply consider the need for every desktop and every user to be able to attach to centralized resources as well as to the Internet.

The factors that can affect the scaling of EIGRP are as follows:

- The amount of information sent between neighbors
- The number of routers that are sent updates
- How far away the routers are that have to send updates
- The number of alternative paths to remote networks

Poorly scaled EIGRP networks can result in the following:

- A route being SIA
- Network congestion
 - Delays
 - Routing information being lost
 - Flapping routes
 - Retransmission
- Router memory running low
- Router CPU overutilized
- Unreliable circuit or unidirectional link

Some of these symptoms are caused by other factors, such as poor design, with resources overwhelmed by the tasks assigned. Often, many of these symptoms are characterized by a route being flagged as SIA, as the router waits for a reply from a neighbor across a network that cannot handle the demands made upon it.

Careful design and placement of network devices can remedy many of the problems seen in a network.

Solutions to EIGRP Scaling Issues

The design of the network is very important to the ability to scale any network. The following solutions revolve around a carefully thought-out network:

- Allocation of addresses should be contiguous to allow summarization.
- A hierarchical tiered network design should be used to allow summarization.
- Sufficient network resources (both hardware and software) on network devices.
- Sufficient bandwidth should be used on WAN links.
- Appropriate EIGRP configuration should be used on WAN links. By default, EIGRP only uses 50 percent of the bandwidth of the link for its traffic. This default may be tuned manually.
- Filters should be used.
- Network monitoring should be used.

EIGRP Design Issues

The major concern in scaling an organizational network is controlling the network overhead that is sent over slow WAN links in particular. The less information about the network, its services, and networks that needs to be sent, the greater the capacity available for the data between clients and servers. Although sending less routing information relieves the network, it gives the routers less information with which to make decisions. Every designer of routing protocols and every network administrator must deal continually with this trade-off. As seen with summarization, static and default routes can lead to poor routing decisions and loss of connectivity.

EIGRP automatically summarizes at the autonomous system boundary and at the classful network boundary. To configure manual configuration, it is first necessary to disable automatic summarization. Summarization is configured at the interface level. This obviously requires careful consideration of the network design in reference to the flow of data and the network topology. Although still a distance vector protocol and proprietary, EIGRP addresses many of the problems related to scaling the network.

Remember that queries must be limited to ensure that EIGRP can properly scale. If queries are allowed to traverse the entire organization, the problems and symptoms described will ravage your network.

Many believe that dividing the organization's network into different EIGRP autonomous systems is a good way of limiting the query range. This is true, because EIGRP does not share updates with

another autonomous systems. However, many organizations that created autonomous systems to replicate OSPF areas naturally redistribute between them so that the entire organization can share routing information. At this point, the query is propagated into the new autonomous system, and the problem continues. Summarization is the best way to limit the query range of EIGRP networks. If a subnet is hidden by summarization, the query will stop at the first router that has no knowledge of it.

Certain topologies, although valid in most instances, pose problems for the EIGRP network. This is true in particular for the hub-and-spoke design often seen implemented between remote sites and regional offices. The popular dual-homed configuration, although providing redundancy, also allows the potential for routers to reflect queries back to one another. Summarization and filters make this network design work well while also allowing queries to be managed effectively.

Foundation Summary

The “Foundation Summary” section of each chapter lists the most important facts from the chapter. Although this section does not list every fact from the chapter that will be on your exam, a well-prepared candidate should, at a minimum, know all the details in each “Foundation Summary” before going to take the exam.

The main concepts of EIGRP are as follows:

- DUAL
- Loop-free networks
- Incremental updates
- Multicast addressing for updates
- Advanced distance vector protocol
- Loop-free routing tables
- Support for different topologies
- Rapid convergence
- Reduced bandwidth use
- Protocol independence at Layer 3
- Compatibility with IGRP
- Easy configuration
- Use of a composite metric
- Unequal-cost load balancing

Cisco identifies four main components of EIGRP:

- Protocol-dependent modules
- RTP
- Neighbor discovery and recovery
- DUAL

Table 13-5 summarizes the EIGRP packet types sent between neighbors.

Table 13-5 *Summary of Packet Types*

Packet Type	Address	Reliable	Unreliable	Purpose
Hello	Multicast		X	To find and maintain neighbors for the neighbor table. The packet has a 0 in the Acknowledgment field.
ACK	Unicast	X		A Hello packet with no data. It has a positive number in the Acknowledgment field.
Update	Unicast and multicast (Reply to a single router is unicast, but a change in topology table is multicast.)	X		Route information sent to affected routers.
Query	Multicast	X		A part of DUAL. Queries are sent out when a route in the topology table goes down and there is no FS.

Figure 13-7 shows the actions taken when a router receives a query from another router asking for an alternative route to a destination. Note that if the queried router has no route to offer, it is still obliged to respond to the querying router.

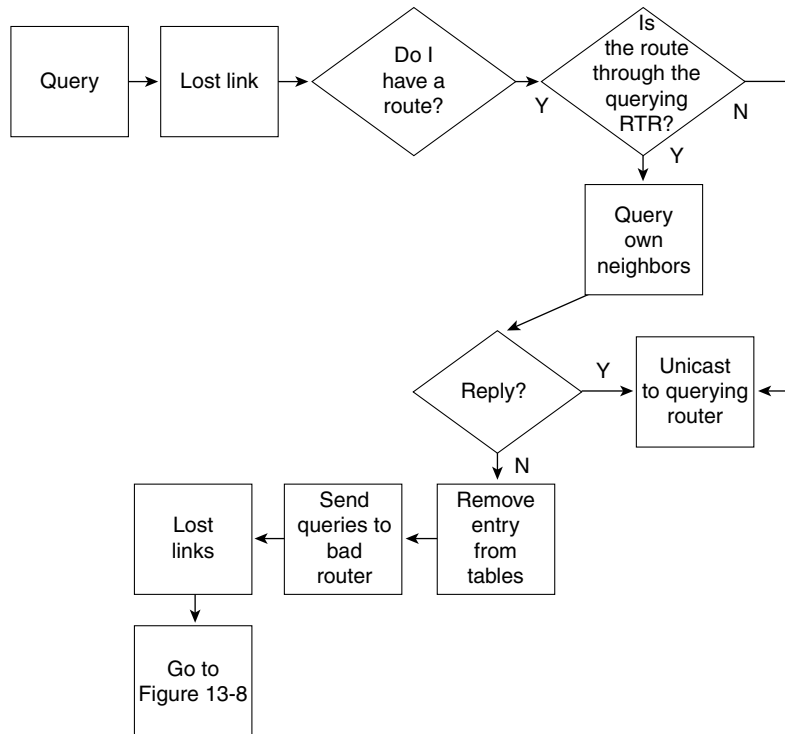
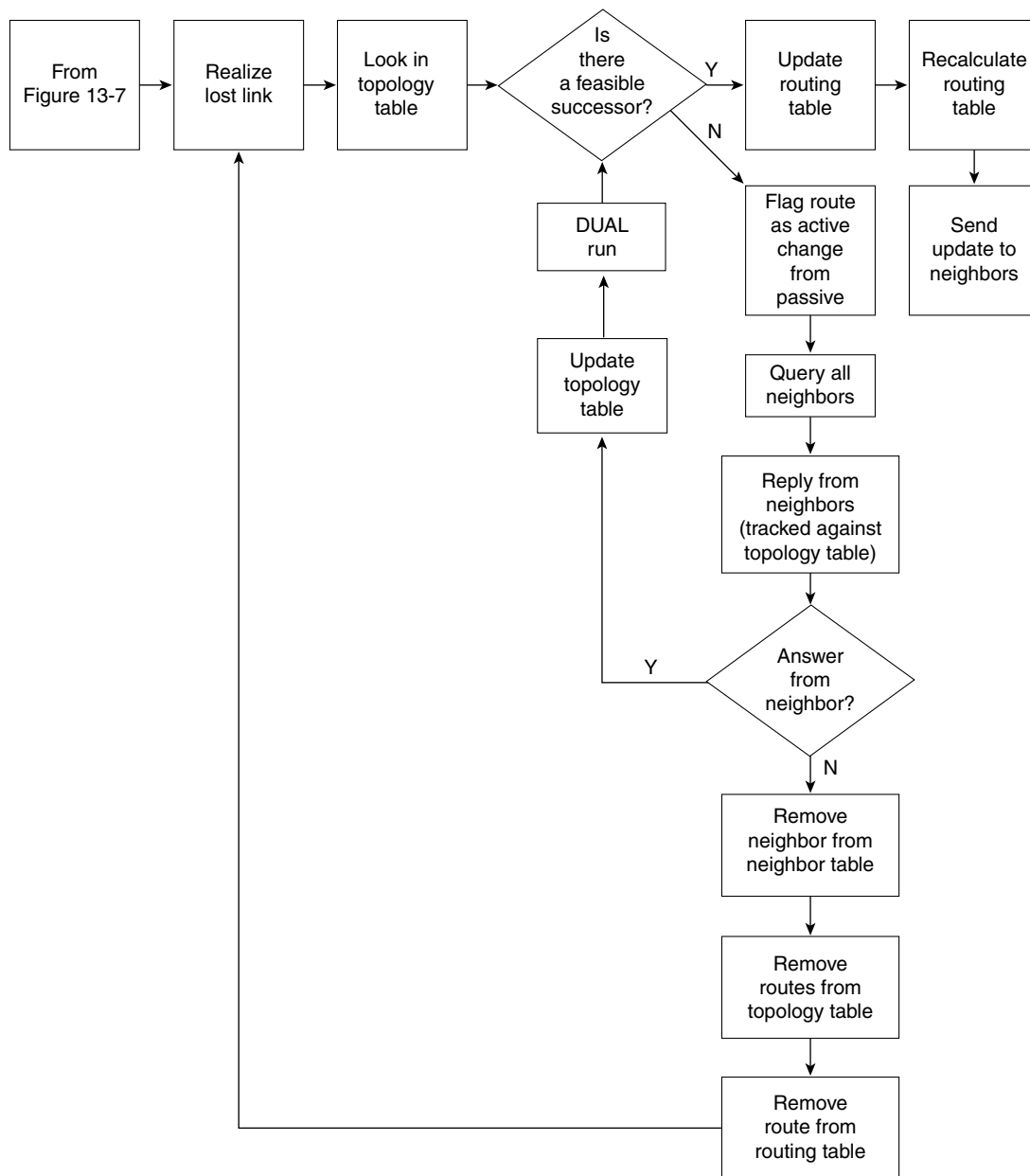
Figure 13-7 *EIGRP—Maintaining the Topology Table, Router D*

Figure 13-8 illustrates the logic flow in a router that realizes a link has been lost, which may occur because a directly connected interface has lost a carrier signal or because the router has received an update or query.

Figure 13-8 *EIGRP—Maintaining the Topology Table, Choosing a Feasible Successor*

Q&A

As mentioned in the introduction, “All About the CCNP, CCDP, and CCIP Certifications,” you have two choices for review questions. The questions that follow next give you a bigger challenge than the exam itself by using an open-ended question format. By reviewing now with this more difficult question format, you can exercise your memory better and prove your conceptual and factual knowledge of this chapter. The answers to these questions are found in Appendix A.

For more practice with examlike question formats, including questions using a router simulator and multichoice questions, use the exam engine on the CD-ROM.

1. If a router does not have a feasible successor, what action will it take?
2. When does EIGRP need to be manually redistributed into another EIGRP process?
3. Which timers are tracked in the neighbor table?
4. What is the difference between an update and a query?
5. When does EIGRP recalculate the topology table?
6. EIGRP has a default limit set on the amount of bandwidth that it can use for EIGRP packets. What is the default percentage limit?
7. State two rules for designing a scalable EIGRP network.
8. EIGRP can be used to send information about which three routed protocols?
9. Which EIGRP packets are sent reliably?
10. In what instances will EIGRP automatically redistribute?
11. How long is the holdtime, by default?
12. What is an EIGRP topology table, and what does it contain?
13. What is the advertised distance in EIGRP, and how is it distinguished from the feasible distance?
14. What EIGRP algorithm is run to create entries for the routing table?
15. When does EIGRP place a network in active mode?
16. By default, EIGRP summarizes at which boundary?
17. What is Stuck in Active?
18. State two factors that influence EIGRP scalability.
19. What are reply packets in EIGRP?
20. What conditions must be met for a router to become a neighbor?



This chapter covers the following topics, which you need to understand to pass the CCNP/CCDP/CCIP BSCI exam:

- Configuring EIGRP
- Verifying the EIGRP operation
- Troubleshooting the EIGRP operation

Configuring EIGRP in Enterprise Networks

The topics in this chapter detail the steps to configuring the EIGRP protocol for integrated routing on a Cisco network. This chapter assumes knowledge of routing protocols, in particular, link-state routing protocols and the terminology, concepts, and operation of EIGRP. This chapter introduces EIGRP configuration commands by explaining the required commands and then discussing the optional configuration commands that can be used.

“Do I Know This Already?” Quiz

The purpose of the “Do I Know This Already?” quiz is to help you decide what parts of this chapter to use. If you already intend to read the entire chapter, you do not necessarily need to answer these questions now.

The 12-question quiz, derived from the major sections in the “Foundation Topics” portion of the chapter, helps you to determine how to spend your limited study time.

Table 14-1 outlines the major topics discussed in this chapter and the “Do I Know This Already?” quiz questions that correspond to those topics.

Table 14-1 “Do I Know This Already?” Foundation Topics Section-to-Question Mapping

Foundation Topics Section	Questions Covered in This Section
Configuring EIGRP	1–6
Verifying the EIGRP Operation	7–9
Troubleshooting the EIGRP Operation	10–12

CAUTION The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark this question wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

1. What command starts the EIGRP routing process?
 - a. **router eigrp**
 - b. **router eigrp** *autonomous-system-number*
 - c. **ip routing eigrp**
 - d. **ip eigrp routing**
2. What happens if no **network** command is configured?
 - a. The EIGRP process is not activated on any interface.
 - b. The EIGRP defaults to all interfaces.
 - c. The EIGRP process can receive, but cannot send, updates.
 - d. The router prompts you for the networks.
3. What happens immediately after the **network** command is configured?
 - a. Updates are sent.
 - b. The routing table is created.
 - c. Hellos are sent on appropriate interfaces.
 - d. Networks are advertised.
4. Where in the EIGRP network is it possible to summarize IP addresses?
 - a. At the IANA major network boundary
 - b. At the ASBR
 - c. At the ABR
 - d. Anywhere in the network
5. Where is the command **ip summary-address eigrp** *autonomous-system-number address mask* configured?
 - a. At the routing process
 - b. At the interface

- c. After the **network** command
 - d. At the EXEC command prompt
- 6. What command is used to change the Hello timer?
 - a. **eigrp hello-interval seconds**
 - b. **hello interval seconds**
 - c. **ip hello-interval eigrp *autonomous-system-number* seconds**
 - d. **ip eigrp hello timer seconds**
- 7. Which of the following commands will show the holdtime?
 - a. **show ip eigrp topology**
 - b. **show ip eigrp traffic**
 - c. **show ip eigrp holdtime**
 - d. **show ip eigrp neighbors**
- 8. Which command shows the state of the links with the neighbors?
 - a. **show ip eigrp traffic**
 - b. **show ip eigrp topology**
 - c. **show ip eigrp adjacencies**
 - d. **show ip eigrp database**
- 9. What is shown in the **show ip eigrp traffic** command?
 - a. Updates
 - b. Hellos
 - c. Queries and replies
 - d. ACKs
- 10. What is displayed in the command **debug ip eigrp summary** ?
 - a. A summary of EIGRP activity
 - b. A summary of the contents of the neighbor database
 - c. The process taken when a change is made in a summary route
 - d. A summary of topology database

11. Which command is used to see the types of packets sent and received, as well as the statistics on routing decisions?
- a. **show ip eigrp traffic**
 - b. **show eigrp events**
 - c. **debug ip eigrp packets**
 - d. **debug ip eigrp traffic**
12. How many different packet types can be seen in the command **debug eigrp packet** ?
- a. 7
 - b. 5
 - c. 11
 - d. 7

The answers to the “Do I Know This Already?” quiz are found in Appendix A, “Answers to Chapter ‘Do I Know This Already?’ Quizzes and Q&A Sections.” The suggested choices for your next step are as follows:

- **6 or less overall score** —Read the entire chapter. This includes the “Foundation Topics” and “Foundation Summary” sections, the “Q&A” section, and the “Scenarios” at the end of the chapter.
- **7–9 overall score** —Begin with the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. If you have trouble with these exercises, read the appropriate sections in “Foundation Topics.”
- **10 or more overall score** —If you want more review on these topics, skip to the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. Otherwise, move to the next chapter.

Foundation Topics

Configuring EIGRP

The commands for EIGRP are consistent with the other IP routing protocols. Although IP routing is on automatically, the chosen routing protocol must be configured and the participating interfaces must be identified.

EIGRP allows for variable-length subnet mask (VLSM) and, therefore, summarization, because the mask is sent in the update packets. Although summarization is automatic, EIGRP summarizes at the IANA or major network boundary. To summarize within the IANA number, it must be manually configured. Unlike OSPF, which can summarize only at the area border router (ABR), EIGRP can summarize not only at any router, but also at any interface on any router.

NOTE EIGRP has evolved over the past few years. It is essential that, in a practical situation, you research the commands and configuration for the IOS software code level that is installed in your network.

This section covers the following:

- Required commands for configuring EIGRP
- Optional commands for EIGRP
- Optional EIGRP commands specific to WANs

Required Commands for Configuring EIGRP

The router needs to understand how to participate in the EIGRP network. Therefore, it requires the following:

- **The EIGRP process** —The routing protocol needs to be started on the router.
- **The EIGRP autonomous system number** —All routers sharing routing updates and participating in the larger network must be identified as part of the same autonomous system. A router will not accept an update from a router configured with a different autonomous system number.
- **Participating router interfaces** —The router might not want to have all its interfaces to send or receive EIGRP routing updates. A classic example is a dialup line to a remote office. If there is only one subnet at the remote office, it would be more efficient to use default and static route commands, because any updates would dial the line.

By default (unless the setup script is used), there is no IP routing protocol running on the Cisco router. This is not true of other protocols, however. If an IPX network address is configured on an interface, for example, the IPX RIP process will be automatically started.

To configure EIGRP as the routing protocol, the following command syntax is used:

```
Router(config)#router eigrp autonomous-system-number
```

Although EIGRP has been turned on, it has no information about how to operate. The connected networks that are to be sent in the EIGRP updates and the interfaces that participate in the EIGRP updates must be defined. If the EIGRP information is not specified, the process with insufficient configuration will never start.

NOTE Most versions of the IOS software do not offer an error message when the configuration is incomplete, which can make troubleshooting more difficult. Refer to the section titled “Verifying the EIGRP Operation,” later in this chapter for more information.

The following command syntax shows the use of the **network** command prior to Cisco IOS release 12.0(4)T:

```
Router(config-router)#network network-number
```

The **network** command in EIGRP plays a similar role to that of the **network** command in RIP or IGRP. Like OSPF, in which it is possible to identify the specific address of an interface, the **network** command for EIGRP can be stated with a mask option, allowing you to identify which interfaces are to run EIGRP. However, it is important to remember that EIGRP does not use areas. The ability to define the network mask was introduced in Cisco IOS release 12.04 (T).

NOTE A common error is to configure the **network** command with an inappropriate wildcard mask when you are confused as to which class of address is being used. Be sure to identify the correct wildcard mask to avoid the situation in which EIGRP only runs on some, if any, of the interfaces.

The new syntax is as follows:

```
Router(config-router)#network network-number [wildcard-network-mask]  
Router(config-router)#no network network-number [wildcard-network-mask]
```

The following syntax illustrates the use of the **network** command (the router has two Ethernet interfaces):

```
Router(config)#interface e1  
Router(config-if)#ip address 155.16.1.1 255.255.255.0  
!  
Router(config)#interface e2  
Router(config-if)#ip address 155.16.2.2 255.255.255.0
```

The following command indicates that EIGRP will run on interface e1 only:

```
Router(config)#router eigrp 100  
Router(config-router)#network 155.16.1.1 0.0.0.0
```

In versions prior to the IOS release 12.04 (T), the network command acted differently. As soon as the first part of the command was configured, the operating system corrected the address to the Internet or major (classful) network number. In this case, example network 155.16.1.1 would be connected to 155.16.0.0, which would include both e1 and e2.

After the network has been defined to EIGRP, it identifies the interfaces directly connected to the routers that share that network address. In some instances, it is not a good idea to have EIGRP updates running across certain interfaces, for example, links connecting to stub routers or to another routing protocol or autonomous system. Before 12.04(T), you would prevent EIGRP from sending updates through an interface by issuing the **passive-interface** command.

When a passive interface is created, it prevents Hellos from being sent between routers. This means that the routers cannot become neighbors, which results in no routing updates being either sent or received. However, the address of the interface is sent in updates out of nonpassive interfaces.

The passive interface allows the network address to be connected to stub routers (a typical frame relay hub-and-spoke configuration). The Ethernet network on the other side of the stub router might be included into the routing tables and propagated throughout the network, without using resources on the Ethernet link or on the router.

Once the interfaces on the router that are participating in the EIGRP domain using the **network** command are identified, the following happens:

- Updates are received on the interface.
- Updates are sent out the interfaces.
- The network are advertised out all EIGRP interfaces.
- If appropriate, the Hello protocol is propagated.

Optional Commands for Configuring EIGRP

The optional commands are used to tune the way EIGRP works within your network. They should be used in reference to the design of the network and its technical requirements.

This section considers the following optional EIGRP commands described in Table 14-2.

Table 14-2 *Optional Commands for Configuring EIGRP*

Command	Description
no auto-summary	Turns off auto summary, allowing the configuration of manual summarization.
ip summary-address	Manual configuration of summarization.
eigrp stub	Configures a stub router.
variance	Configures unequal load balancing over multiple routes.
ip hello-interval eigrp <i>autonomous-system-number seconds</i>	Changes the number of seconds between Hellos.
ip hold-time eigrp <i>autonomous-system-number seconds</i>	Changes the length of time before a route is considered dead because the neighbor has not sent a Hello within the required time.
bandwidth	Changes the bandwidth setting on an interface, which affects the EIGRP metric calculation and the amount of EIGRP traffic that is sent through the interface.
bandwidth-percent	Changes the amount of the bandwidth that EIGRP traffic can use. The default is 50%.

Summarization with EIGRP

Summarization in EIGRP solves the same scaling issues seen in other networks. The difference in the configuration between EIGRP and OSPF is that the OSPF is summarized only at the area boundary. Because EIGRP does not use the concept of areas, summarization can be configured on any router interface in the network. Consideration of where to summarize is determined by the hierarchical structure of the network. If summarization is not configured, EIGRP will automatically summarize at the class boundary.

NOTE Other chapters in this book have dealt with summarization in detail. For the sake of brevity, only those details related specifically to EIGRP are conveyed here. For more information about summarization, refer to Chapter 8, “Using OSPF Across Multiple Areas,” and the sections “Design Considerations in Multiple Area OSPF” and “Summarization” in particular.

Summarization has advantages for EIGRP above and beyond the benefit of smaller routing tables, as explained in Chapter 13, “Using EIGRP in Enterprise Networks.” Summarization reduces the amount of resources needed by both the network and the routers within the network. The reduced routing tables speed up the lookup when forwarding data that is process switched. Summarization also reduces the

scope of the queries sent out by a router. If a router has no feasible successor, it queries its neighbor for an alternative route. If the neighbor has no route to offer, the query is forwarded on until a route is found or the search is exhausted. If summarization has been configured, the route that is being queried might have been summarized, and thus the query will end. Thus, summarization can limit the scope of the query, because when a subnet is hidden in summarization, a reply of unknown network will be returned to the router that can purge the route from the databases.

There are two commands for summarization with EIGRP: **no auto-summary** and **ip summary-address eigrp autonomous-system-number address mask**. The first command, **no auto-summary**, disables the automatic summarization. This command applies to the entire router. With **no auto-summary** configured, information on all the known subnets is sent out of every interface. If there are slow serial interfaces or congested links, these links could become overwhelmed. The solution is to configure the **ip summary-address eigrp** command on all interfaces, which in turn demands careful deployment of addresses.

Manual summarization is configured at the interface level, as shown here:

```
Router(config)#interface S0
Router(config-if)#ip summary-address eigrp autonomous-system-number address mask
```

Stub Routers

IOS software release 12.0 made it possible for you to configure a remote router as an EIGRP stub router. A stub router is typically used on small capacity routers in a hub-and-spoke WAN environment. The stub router in EIGRP is similar to the concept of On Demand Routing (ODR) described in Chapter 1, “IP Routing Principles.” ODR is used in similar situations but has no routing protocol configured on the stub router. ODR uses CDP to maintain connectivity between the stub routers and core router sending a default route to the stub. Stub routers in EIGRP networks use EIGRP to send limited information between the stub and the core routers.

As in ODR, the router in an EIGRP network has no other neighbors and accesses the network through a distribution layer router. It is not necessary, therefore, for this remote router to have a complete routing table that may overwhelm its limited resources. The remote router needs only a default route to the distribution router that can serve all its needs.

Another reason to configure the remote router as a stub is to lend a hand to the rest of the network. If a query is sent to a remote router, the delays involved can result in the path being Stuck in Active (SIA). If the stub configuration has been applied, the router responds to queries as inaccessible, thus limiting the scope of the query range and preventing SIA from occurring.

The following command structure shows the syntax of the **eigrp stub** command:

```
Router(config-router)# eigrp stub [receive-only | connected | static | summary]
```

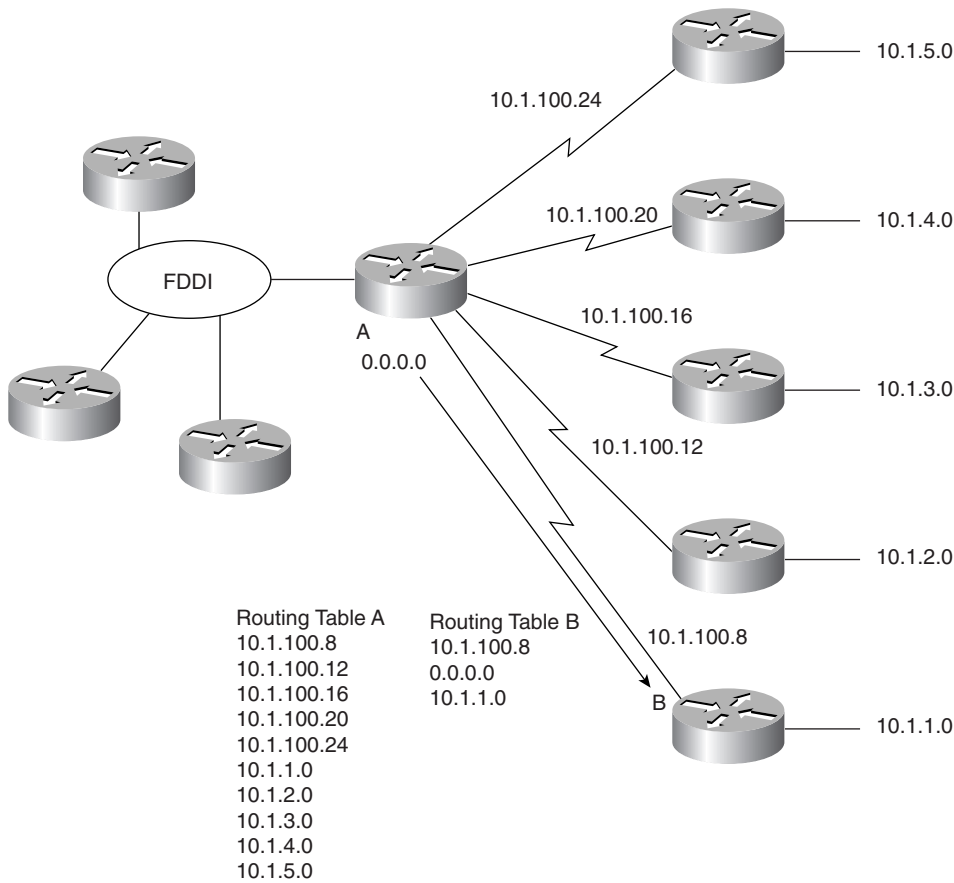
Table 14-3 explains the syntax of this command.

Table 14-3 The **eigrp stub** Command Syntax and Description

Parameter	Description
receive-only	(Optional) Sets the router as a receive-only neighbor
connected	(Optional) Advertises connected routes
static	(Optional) Advertises static routes
summary	(Optional) Advertises summary routes

Figure 14-1 shows a group of routers connected over WAN links. These routers are stub routers because they have no other networks connected to them.

Figure 14-1 The **eigrp stub** Router Command



Example 14-1 is the configuration for Router B in Figure 14-1.

Example 14-1 *The EIGRP Stub Router Command*

```
RouterB(config)#router eigrp 100
RouterB(config-router)#network 10.0.0.0 255.0.0.0
RouterB(config-router)#eigrp stub
```

Load Balancing in EIGRP

EIGRP automatically load balances across links of equal cost. Whether the traffic is sent on a per-destination or round-robin basis depends on the internal switching within the router. It is possible to configure EIGRP to load balance across unequal-cost paths using the **variance** command.

The **variance** command allows the administrator to identify the metric scope for including additional paths by the use of a multiplier parameter. The command structure follows:

```
Router(config-router)#variance multiplier
```

The *multiplier* argument is the metric value used for load balancing. It can be a value from 1 to 128. The default is 1, which means equal-cost load balancing.

Example 14-2 shows the configuration of the **variance** command.

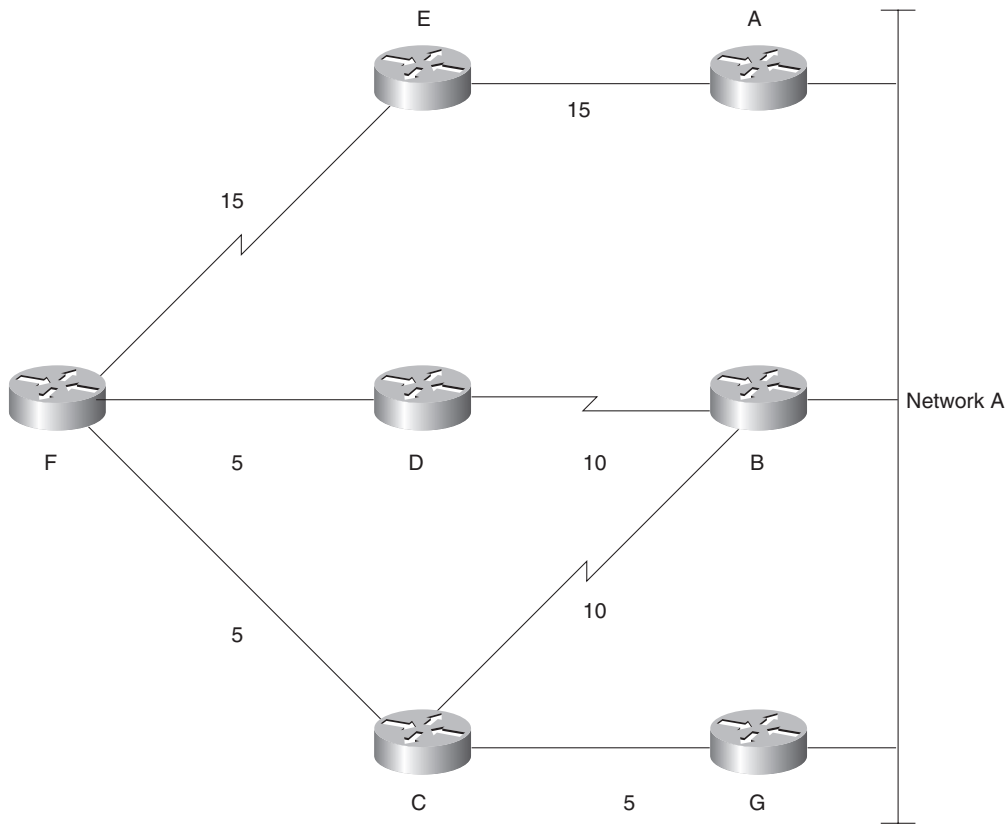
Example 14-2 *The variance Command*

```
RouterB(config)#router eigrp 100
RouterB(config-router)#network 10.0.0.0 0.0.0.255
RouterB(config-router)#variance 2
```

If the variance number is higher than the default of 1, the EIGRP process multiplies the best (lowest) cost or metric value for a path by the number stated as the variance multiplier. All paths to the same destination that have metrics within this new range are now included in load balancing. The amount of traffic sent over each link is proportional to the metric for the path.

For example, the route to Network A in Figure 14-2 has four paths to it from Router F, and the best path gave a metric value of 10. The available routes shown in Figure 14-2 reflect these paths:

- F to E to A = 30
- F to D to B = 15
- F to C to B = 15
- F to C to G = 10

Figure 14-2 *Including Unequal Paths in Load Balancing*

NOTE Only those paths that are in the topology table as feasible successor (FS) are eligible to be included in the **variance** command. Also, the example and figure shown are highly simplified for the purpose of explanation.

If the **variance** command was configured with a variance, or multiplier, of 2, the best metric is $10 * 2 = 20$. Any route with a metric of 20 or better will be placed in the routing table.

These paths would all load-balance traffic from Router F to Network A:

F to D to B = 15
 F to C to B = 15
 F to C to G = 10

One-and-a-half packets would be sent across the path F to C to G for every one packet sent across the other two available paths.

The router rounds the number of packets to be sent to 2 packets, giving a traffic ratio of 3:2.

Tuning the EIGRP Process

There are many ways to tune a network, including load balancing across multiple paths, summarizing routes, and reducing the frequency of the update timers. There is, however, a trade-off between reducing the resources required to maintain the network and the stability of the network. The fewer Hellos that are sent out, for instance, the longer the network might take to notice a failure, and convergence of the network would be delayed. When the network does not have an accurate understanding of the available routes, the router cannot forward packets with any confidence.

The Hello timer and the receipt of ACKs are particularly important because EIGRP sends out incremental updates. The process sends updates only when a failure is seen or to advertise a new network, which makes it important to have a reliable and immediate method of noticing the link has died. Hence, reliable transport protocol (RTP) for EIGRP was created. Furthermore, it is the responsibility of the neighbor to first notice and then inform the rest of the network through an update that the network is no longer available.

You can configure the Hello timer, but you must consider how changing such a fundamental element impacts the accurate running of EIGRP. The hold timer indicates how long a route is held without a Hello being heard before the route is deemed to be no longer in existence. Both the Hello timer and the hold timer are discussed in the next sections.

The Hello Interval Timer

Tuning the Hello interval directly affects the ability of the network to notice a change in the state of a neighbor. Only after a router's interface is recognized as being down, or the router has failed to hear from a neighbor after a proscribed amount of time, does the router declare the neighbor dead and take the necessary action to update the routing table and the rest of the network.

For these reasons, the **ip hello-interval eigrp** command is typically used to decrease the time between Hellos to ensure that the network is more stable and converges more quickly. Although this increases the amount of bandwidth consumed, it is a minimal cost. This command becomes very useful in WANs, particularly when nonbroadcast multiaccess (NBMA) clouds are used. EIGRP treats both Frame Relay and Switched Multimegabit Data Service (SMDS) as NBMA technologies, resulting in Hello timers that assume a low bandwidth medium (less than T1 speeds) and that set the timer to 60 seconds by default.

The command to change how often the Hellos are sent to neighbors is as follows:

```
Router(config-if)#ip hello-interval eigrp autonomous-system-number seconds
```

The autonomous system number identifies the EIGRP process to the autonomous system.

The number of seconds to wait between each Hello is configured at the end of the command. An example of this configuration follows:

```
Router(config)#interface Serial 0
Router(config-if)#ip hello-interval eigrp 100 10
```

The defaults for Hello packet timers are as follows:

- High bandwidth links (every 5 seconds):
 - Broadcast media, such as Ethernet, Token Ring, and FDDI
 - Point-to-point serial links, such as PPP or HDLC leased circuits, Frame Relay point-to-point subinterfaces, and ATM
 - Point-to-point subinterfaces
 - High bandwidth (greater than T1) multipoint circuits, such as ISDN PRI and Frame Relay
- Low bandwidth links (every 60 seconds):
 - Multipoint circuits T1 bandwidth or slower, such as Frame Relay multipoint interfaces, ATM multipoint interfaces, and ATM
 - Switched virtual circuits and ISDN BRI

The Hold Timer

The *holdtime* is how long the router waits without hearing a Hello from the neighbor before pronouncing it unavailable. The holdtime is three times that of the Hello timer by default, but changing the rate at which EIGRP sends Hello packets does not automatically change the holdtime. The hold timer must be changed manually using the **ip hold-time eigrp** command. The command syntax follows:

```
Router(config-if)#ip hold-time eigrp autonomous-system-number seconds
```

The following example shows the syntax in context:

```
Router(config)#interface ethernet 0
Router(config-if)#ip hold-time eigrp 100 30
```

Optional EIGRP Commands over WANs

There are always particular design and configuration issues concerning WANs. With WANs, more than at any other point in the network, you are likely to deal with limited resources. Therefore, it is with WAN topologies that you will use the **bandwidth** and **bandwidth-percent** commands, because they determine the link resources allocated to EIGRP updates and are used to calculate the metrics assigned to routes.

A perennial concern of network administrators is the amount of bandwidth used for overhead traffic. Administrators want to minimize the amount of network control traffic sent through the network to maximize the bandwidth available for user data. One of the major benefits of both EIGRP and OSPF is that they send as little network traffic as possible. This has the advantages of decreasing the convergence time of the network and ensuring that the network traffic that is sent arrives at the destination.

EIGRP Defaults in Bandwidth Utilization

EIGRP will not use more than 50 percent of the stated bandwidth on a link for its own routing traffic. The **bandwidth** command used on the interfaces of a Cisco router allows the default settings on links to work as intended, by stating the actual bandwidth of the link. This is often necessary on serial links because the default bandwidth is 1.544 Mbps or a T1. If in reality the link is 56 kbps, it is easy to see how EIGRP could saturate the link. EIGRP allows itself to use up to 50 percent of a T1 link (772 kbps), far exceeding the real bandwidth of the line. This could mean not only dropping data packets because of congestion but also dropping EIGRP packets. This will cause confusion in the network, not to mention miscalculated routes, retransmission, and user frustration as the network slows.

Other technologies, such as OSPF and SMDS on a Cisco router use the bandwidth value to make decisions. You need to ensure that the bandwidth stated is indeed the speed of the link. When you issue the **show interface** command, the configured bandwidth of the link will be shown along with a field identifying the load on the line. The load is the amount of traffic sent out of the interface, proportional to the bandwidth of the link, in which the bandwidth is the stated bandwidth and not the actual speed of the physical interface.

NOTE If it is necessary to artificially lower the bandwidth using the **bandwidth** command, this should be done in consideration of the other network applications.

The bandwidth is a logical construct whose value can have wide-reaching implications on the function of your network. It does not affect the actual speed of the link. In fact, it is practical to configure the **bandwidth** command only on serial lines, where the speed of the link will vary considerably. The following section provides further guidelines from Cisco on bandwidth configuration.

Rules in Configuring Bandwidth over an NBMA Cloud

EIGRP works well over all WAN environments, including point-to-point and NBMA such as Frame Relay, X25, or ATM. The NBMA topology can include either point-to-point subinterfaces or an NBMA hybrid, which is a combination of point-to-point and multipoint configurations.

Cisco identifies three rules that you should follow when configuring EIGRP over an NBMA cloud:

- EIGRP traffic should not exceed the committed information rate (CIR) capacity of the virtual circuit (VC).
- EIGRP's aggregated traffic over all the VCs should not exceed the access line speed of the interface.
- The bandwidth allocated to EIGRP on each VC must be the same in both directions.

If you understand and follow these rules, EIGRP works well over the WAN. If you do not take care in the configuration of the WAN, EIGRP can swamp the network.

Configuring Bandwidth over a Multipoint Network

The configuration of the **bandwidth** command in an NBMA cloud depends on the design of the VCs. If the serial line has many VCs in a multipoint configuration, EIGRP will evenly distribute its overhead between the VCs, without the use of subinterfaces. The **bandwidth** command should therefore reflect the access link speed into the Frame Relay cloud. If the serial interface is accessing an NBMA environment such as Frame Relay, the situation is straightforward. Your company might have five VCs from your router's serial interface, each carrying 56 kbps. The access link will need a capacity of $5 * 56$ kbps. Remember, the aggregate configured bandwidth cannot exceed the access speed of the interface.

Configuring Bandwidth over a Hybrid Multipoint Network

If the multipoint network has differing speeds allocated to the VCs, a more complex solution is needed. There are two main approaches:

- Take the lowest CIR and simply multiply it by the number of circuits. This is applied to the physical interface. The problem with this configuration is that the higher-bandwidth links will be underutilized for some things.
- If possible, it is much easier to configure and manage an environment that has used subinterfaces, where a VC is logically treated as if it were a separate interface or point-to-point. The **bandwidth** command can be configured on each subinterface, which allows different speeds on each VC. In this solution, subinterfaces are configured for the links with the differing CIRs. The links that have the same configured CIR are presented as a single subinterface with a bandwidth, which reflects the aggregate CIR of all the circuits.

Cisco recommends this as the preferred solution.

The following syntax shows the structure of the **bandwidth** command:

```
Router#interface S0
Router(config-if)#bandwidth speed-of-line
```

Configuring the Pure Point-to-Point Network

If there are many VCs, there might not be enough bandwidth at the access speed of the interface to support the aggregate EIGRP traffic. The subinterfaces should be configured with a bandwidth that is much lower than the real speed of the circuit. In this case, it is necessary to use the **bandwidth-percent** command to indicate to the EIGRP process that it can still function.

As you learned in the previous section, EIGRP limits itself to 50 percent of the value specified in the **bandwidth** command, or if the bandwidth command is not set, the interface defaults. If you need to limit this percentage further, the upper limit that EIGRP uses can be stated as a percentage of the **bandwidth** command.

The **ip bandwidth-percent eigrp** command interacts with the **bandwidth** command on the interface. You would use this command primarily because in your network, the **bandwidth** command does not reflect the true speed of the link. The **bandwidth** command might have been altered to manipulate the routing metric and path selection of a routing protocol, such as IGRP or OSPF. It might be better to use other methods of controlling the routing metric and return the bandwidth to a true value. Otherwise, the **ip bandwidth-percent eigrp** command is available. It is possible to set a bandwidth percent that is larger than the stated bandwidth. This is with the understanding that, although the bandwidth might be stated to be 56 kbps, the link is in fact 256 kbps. The following shows the structure of the **ip bandwidth-percent eigrp** command:

```
Router(config)#interface S0
Router(config-if)#ip bandwidth-percent eigrp autonomous-system-number percent
```

Verifying the EIGRP Operation

Understanding the output of the commands discussed in this section is important, not just because they might constitute questions on the exam, but because they reflect your conceptual understanding of the subject. The ability to analyze what is happening on the network demands a thorough understanding of the concepts explained in this chapter. This skill is required in interpreting the output of a **show** command.

The ability to interpret these **show** command output examples in conjunction with the physical and logical topology diagrams of your organization will ensure your understanding of the operation of EIGRP.

This section deals with the **show** commands shown in Table 14-4.

Table 14-4 *EIGRP show Commands*

Command Option	Description
show ip eigrp neighbors	Gives detailed information about the neighbors. This command records the communication between the router and the neighbors in addition to the interface and address by which they communicate.
show ip eigrp topology	Gives details about the routes held in the topology table, detailed information on the networks that the router is aware of and the preferred paths to those networks, and the next logical hop as the first step in the path. The router will track the EIGRP packets that have been sent to neighbors in this table.
show ip eigrp topology all-links	Gives details about all the routes and alternative paths held in the topology table. The router will track the EIGRP packets that have been sent to neighbors in this table.
show ip eigrp traffic	Gives information about the aggregate traffic sent to and from the EIGRP process.

The EIGRP **show** commands are highly detailed and give a comprehensive understanding of the state of the network. The other commands generic to IP—**show ip route** and **show ip protocols** , as described in Chapter 7, “Configuring OSPF in a Single Area”—are also useful in the maintenance of EIGRP.

The show ip eigrp neighbors Command

This **show ip eigrp neighbors** command shows the neighbor table. The syntax is as follows:

```
Router#show ip eigrp neighbors [type number]
```

Example 14-3 shows the output of this command.

Example 14-1 *The show ip eigrp neighbors Output*

Router# show ip eigrp neighbors							
IP-EIGRP Neighbors for process 100							
Address	interface	Holdtime	Uptime	Q	Seq	SRTT	RTO
		(secs)	(h:m:s)	Count	Num	(ms)	(ms)
140.100.48.22	Ethernet1	13	0:00:41	0	11	4	20
140.100.32.22	Ethernet0	14	0:02:01	0	10	12	24
140.100.32.31	Ethernet0	12	0:02:02	0	4	5	2

Table 14-5 explains the meaning of the important fields in Example 14-3.

Table 14-5 *Explanation of the show ip eigrp neighbors Command Results*

Field	Explanation
Process 100	The autonomous system number used to identify routers from whom to accept routing updates.
Address	IP address of the EIGRP neighbor.
Interface	Interface on which the router is receiving Hello packets from the neighbor.
Holdtime	Length of time, in seconds, that the router will wait to hear from the neighbor before declaring it down. The default is 15 seconds.
Uptime	Time—measured in hours, minutes, and seconds—since the router first heard from this neighbor.
Q Count	Number of EIGRP packets (update, query, and reply) that the router has queued and is waiting to send.
Seq Num	The sequence number of the last packet that was received from the neighbor.
SRTT	Smooth round-trip time. The time is measured in milliseconds and is measured from the sending of the packet to the receipt of an acknowledgment from the neighbor.
RTO	Retransmission timeout, in milliseconds. This shows how long the router will wait for an acknowledgment before it retransmits the packet.

The show ip eigrp topology Command

The **show ip eigrp topology** command shows the topology table. This command allows for the analysis of DUAL. It shows whether the successor or the route is in an active or passive state. The syntax is as follows:

```
Router#show ip eigrp topology [autonomous-system-number | [[ip-address] mask]]
```

Example 14-4 shows the output of this command.

Example 14-1 *The show ip eigrp topology Output*

```
Router# show ip eigrp topology
IP-EIGRP Topology Table for process 100
Codes:P - Passive, A - Active, U - Update, Q - Query, R - Reply, r - Reply status
P 140.100.56.0 255.255.255.0, 2 successors, FD is 0
  via 140.100.32.22 (46251776/46226176), Ethernet0
  via 140.100.48.22 (46251776/46226176), Ethernet1
  via 140.100.32.31 (46277376/46251776), Ethernet0
P 140.100.48.0 255.255.255.0, 1 successors, FD is 307200
```

continues

Example 14-1 The show ip eigrp topology Output (Continued)

```
via Connected, Ethernet1
via 140.100.48.22 (307200/281600), Ethernet1
140.100.32.22 (307200/281600), Ethernet0
via 140.100.32.31 (332800/307200), Ethernet0
```

Table 14-6 explains the meaning of the important fields in Example 14-4.

Table 14-6 Explanation of the show ip eigrp topology Command Results

Field	Explanation
P	Passive—The router has not received any EIGRP input from a neighbor, and the network is assumed to be stable.
A	Active—When a route or successor is down, the router attempts to find an alternative path. After local computation, the router realizes that it must query the neighbor to see whether it can find a feasible successor or path.
U	Update—A value in this field identifies that the router has sent an update packet to a neighbor.
Q	Query—A value in this field identifies that the router has sent a query packet to a neighbor.
R	Reply—A value here shows that the router has sent a reply to the neighbor.
r	This is used in conjunction with the query counter; the router has sent out a query and is awaiting a reply.
140.100.48.0	This is the destination IP network number.
255.255.255.0	This is the destination subnet mask.
successors	This is the number of routes or the next logical hop. The number stated here is the same as the number of applicable routes in the routing table.
FD	Feasible distance—This is the metric or cost to the destination from the router.
Replies	This is the number of replies that the router is still waiting for from this neighbor. This is relevant only when the route is in an active state and is therefore not shown in Example 14-4.
State	This is the EIGRP state of the route. It can be the number 0, 1, 2, or 3. This is relevant when the destination is active and is therefore not shown in Example 14-4.
via	This is the address of the next logical hop, or the neighbor that told the router about this route. The first Ns of these entries are the current successors. The remaining entries on the list are feasible successors.

Table 14-6 *Explanation of the show ip eigrp topology Command Results (Continued)*

Field	Explanation
(46251776/46226176)	The first number is the EIGRP metric that represents the feasible distance, or the cost to the destination. The number after the slash is the EIGRP metric that the peer advertised, or the advertised distance.
Ethernet0	This is the interface through which the EIGRP packets were received and, therefore, is the outgoing interface.

The show ip eigrp traffic Command

The **show ip eigrp traffic** command shows the EIGRP traffic received and generated by the router. The following is the command syntax:

```
Router#show ip eigrp traffic [autonomous-system-number]
```

Example 14-5 shows the output of this command.

Example 14-1 *The show ip eigrp traffic Command Output*

Router# show ip eigrp traffic
IP-EIGRP Traffic Statistics for process 100
Hellos sent/received: 218/205
Updates sent/received: 7/23
Queries sent/received: 2/0
Replies sent/received: 0/2
Acks sent/received: 21/14

Table 14-7 explains the meaning of the important fields in Example 14-5.

Table 14-7 *Explanation of the show ip eigrp traffic Command Output*

Field	Explanation
process 100	The autonomous system number, used to identify routers from whom to accept routing updates
Hellos sent/received	Number of Hello packets sent and received by the router
Updates sent/received	Number of update packets sent and received by the router
Queries sent/received	Number of query packets sent and received by the router
Replies sent/received	Number of reply packets sent and received by the router
Acks sent/received	Number of acknowledgment packets sent and received by the router

Troubleshooting the EIGRP Operation

Many methods and tools help in troubleshooting any network. One of the main keys is documentation, for several reasons: For example, you can see progress and easily eliminate the obvious in a checklist manner, and you can clearly explain the problem and the steps taken so far in solving it if you need to call in expert help. Cisco provides many tools both on its web page and in service contracts to help solve your network problems. One of the mainstays in troubleshooting any routing protocol is the group of **debug** commands, which provide the ability to see traffic and router processes in real time.

Care should be exercised in the use of the **debug** command, because it can be very greedy in terms of the resources that it consumes. It should be used only for a specific option and for a finite time.

The options available for monitoring EIGRP are covered in Table 14-8.

Table 14-8 *The debug Command Options for EIGRP*

Command Option	Description
debug eigrp packet	Shows the packets sent and received by the router. The packet types to be monitored can be selected. Up to 11 types are available.
debug eigrp neighbors	Shows the Hello packets sent and received by the router and the neighbors discovered by this process.
debug ip eigrp route	The default if the command debug ip eigrp is issued. Shows dynamic changes made to the routing table.
debug ip eigrp summary	Shows the process taken when a summary (manual or auto) is changed on the router.

Foundation Summary

The “Foundation Summary” section of each chapter lists the most important facts from the chapter. Although this section does not list every fact from the chapter that will be on your exam, a well-prepared candidate should, at a minimum, know all the details in each “Foundation Summary” before going to take the exam.

Table 14-9 summarizes the commands covered in this chapter.

Table 14-9 *Summary of Commands*

Command	Function
Router(config)# router eigrp <i>autonomous-system-number</i>	Starts the EIGRP processes on the router with the specified autonomous system number.
Router(config-router)# network <i>network-number wildcard-mask</i>	Shows the networks to be advertised.
Router(config-router)# no auto-summary	Given a hierarchical addressing design, disables the automatic summarization to the Internet IANA network address.
Router(config-if)# ip summary-address eigrp <i>autonomous-system-number address mask</i>	Enables you to summarize the networks manually, having disabled automatic summarization.
Router(config-router)# eigrp stub [receive-only connected static summary]	Used on small-capacity routers in a hub-and-spoke WAN environment. The router has no other neighbors and accesses the network through a distribution layer router. Only a small routing table is required at the stub router, which is provided with a default route to the advertising ABR.
Router(config-if)# bandwidth <i>speed-of-line</i>	Issued at the interface level and a logical construct to determine the real bandwidth manually. This command is used mainly on serial lines. Bandwidth will influence some routing decisions and dial-on-demand implementations.
Router(config-if)# ip bandwidth-percent eigrp <i>autonomous-system-number percent</i>	Enables you to change the bandwidth percentage. EIGRP, by default, will only take up to 50% of bandwidth.
Router(config-router)# variance <i>multiplier-number</i>	Allows unequal-cost paths to load balance. Paths included in the equation will send a proportional amount of traffic across the unequal links.

(continues)

Table 14-9 *Summary of Commands (Continued)*

Command	Function
Router(config-if)# ip hello-interval eigrp <i>autonomous-system-number seconds</i>	Changes how often the Hello is sent out of an interface. This is used to tune the network by increasing the frequency and thus reducing convergence timers.
Router(config-if)# ip hold-time eigrp <i>autonomous-system-number seconds</i>	This defines how long the router believes its neighbor to be alive without hearing a Hello from the neighbor. By default, it is three times that of the Hello timer. Remember that if the Hello timer is changed, the hold timer is not automatically changed, but must be changed manually.
Router# show ip eigrp neighbors	Displays information drawn from the neighbor table.
Router# show ip eigrp topology	Displays information drawn from the topology table.
Router# show ip eigrp traffic	Shows the EIGRP traffic passing through the router.
Router# debug eigrp packet	Shows the packets sent and received by the router. The packet types to be monitored can be selected. Up to 11 types are available.
Router# debug eigrp neighbors	Shows the Hello packets sent and received by the router and the neighbors discovered by this process.
Router# debug ip eigrp route	The default if the command debug ip eigrp is issued. Shows dynamic changes made to the routing table.
Router# debug ip eigrp summary	Shows the process taken when a summary (manual or auto) is changed on the router.

Q&A

As mentioned in the introduction, “All About the CCNP, CCDP, and CCIP Certifications,” you have two choices for review questions. The questions that follow next give you a bigger challenge than the exam itself by using an open-ended question format. By reviewing now with this more difficult question format, you can exercise your memory better and prove your conceptual and factual knowledge of this chapter. The answers to these questions are found in Appendix A.

For more practice with examlike question formats, including questions using a router simulator and multichoice questions, use the exam engine on the CD-ROM.

1. Which command has superseded the **passive-interface** command for preventing EIGRP traffic from traversing a link?
2. What is the preferred configuration for a hybrid multipoint NBMA network when one VC has a CIR of 56 kbps and the other five VCs each have a CIR of 256 kbps?
3. With four Frame Relay circuits in a multipoint solution and a bandwidth configuration of 224, what is the EIGRP bandwidth allocation per circuit, and where would the **bandwidth** command be configured?
4. Explain the purpose of the command **no auto-summary** .
5. Explain the meaning of the command **ip bandwidth-percent eigrp 63 100** .
6. In what instances will EIGRP automatically redistribute?
7. How long is the holdtime, by default?
8. For what is the **variance** command used?
9. What command is used to display the passive and active state of the routes?
10. What command is used in EIGRP to perform manual summarization?
11. For Frame Relay, when would you configure the physical interface (as opposed to a subinterface) with the **bandwidth** command?
12. Which command is used to display all types of EIGRP packets that are both received and sent by a router?
13. What problems can be solved with the configuration of summarization?
14. Why would you configure an EIGRP router as a stub router?
15. Explain the parameters **receive-only** , **connected** , **static**, and **summary** used in the command **eigrp stub** .

16. When configuring the **variance** command, which routes can be used?
17. Give two reasons why you might wish to change the Hello timer.
18. What should be considered in terms of EIGRP configuration when there are many VCs entering a WAN and no subinterfaces have been configured?
19. On what occasions should you consider configuring the bandwidth on subinterfaces?

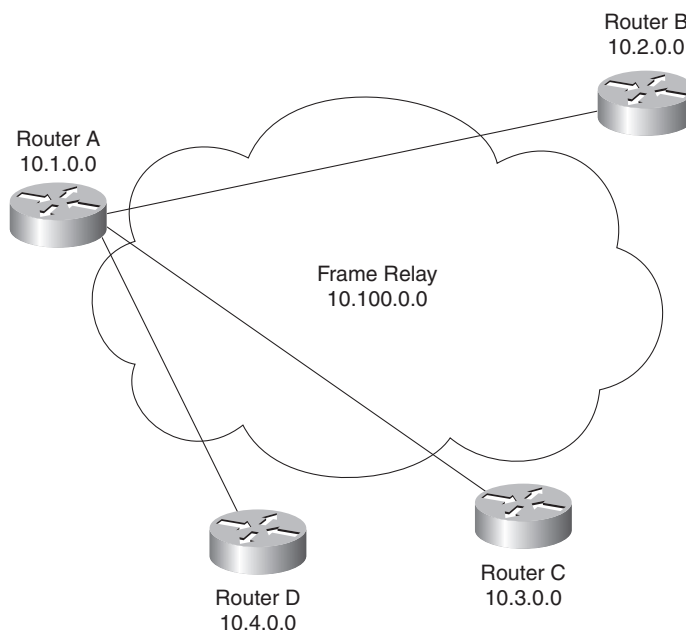
Scenarios

The following scenarios and questions are designed to draw together the content of the chapter and to exercise your understanding of the concepts. There is not necessarily a right answer. The thought process and practice in manipulating the concepts are the goals of this section. The answers to the scenario questions are found at the end of this chapter. The information used in these scenarios was adapted from the Cisco web page, “Cisco Configuration Guidelines.”

Scenario 14-1

The multinational company Gargantuan, Inc., has had a consultant completely readdress the company. The company used the private network 10.0.0.0 and created a hierarchical addressing structure. Refer to Figure 14-3 to see this addressing scheme.

Figure 14-3 *Diagram for Scenario 14-1*



The addressing of the network was a major project, with all the necessary pitfalls that attend such a large exercise. The network is now stable, and it is time to solve the problems that are being experienced as timeouts and network crashes.

The consultant assured the company that the resolution to the delays was the addressing scheme, but although the network is easier to manage, there has been no change in the congestion on the network. In addition, EIGRP appears to be losing routes from its routing tables, which is adding to the problem.

The consultant was correct: The network needed to be readdressed to allow EIGRP to function effectively. Unfortunately, the company did not read the report carefully and missed the other part of the solution.

1. What needs to be done in addition to solve the addressing problems that make EIGRP inefficient? Give the configuration commands necessary to activate this solution on Router A.

The WAN is a Frame Relay cloud, and Router A is the hub in the hub-and-spoke configuration. Each VC is 56 kbps.

2. Give the commands to configure Router A for EIGRP over this NBMA cloud.
3. Give the commands to configure Router B for EIGRP over this NBMA cloud.

Scenario 14-2

Given the configuration of EIGRP in Example 14-6, perform the tasks and answer the questions listed. The WAN has light user traffic and is a hub-and-spoke configuration, as shown in Figure 14-4.

Example 14-1 Scenario 14-2 Configuration for Router A

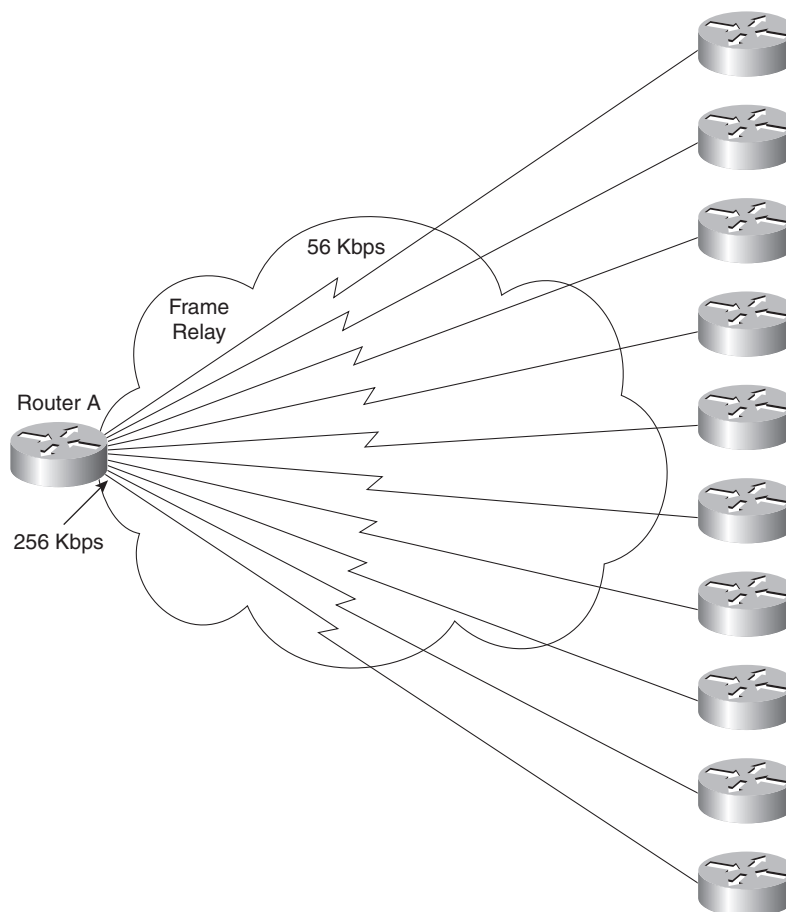
```
Router#show config
interface Serial 0
  encapsulation frame-relay

  interface Serial 0.1 point-to-point
  bandwidth 25
  ip bandwidth-percent eigrp 123 90

  interface Serial 0.2 point-to-point
  bandwidth 25
  ip bandwidth-percent eigrp 123 90

  ...
```

The 256-kbps access line to the hub has 56-kbps access lines to each of ten spoke sites. Each link has a Frame Relay CIR of 56 kbps. The access line to each router reflects the CIR. The access line to the hub router, Router A, is 256 kbps, but the CIR of the hub is the same as its access line.

Figure 14-4 *Diagram for Scenario 14-2*

From a Frame Relay perspective, a circuit is considered oversubscribed when the sum of CIRs of the remote circuits is higher than the CIR of the hub location. With ten links, each with a CIR of 56 kbps, this circuit is clearly oversubscribed.

1. How much bandwidth has each circuit been allocated? Why was this value chosen by the administrator?
2. What bandwidth utilization is available to EIGRP? Why was this value chosen by the administrator?
3. If Router A fails, what would the effect be on the network?
4. Is summarization possible only on the routers entering the WAN cloud, or is it possible on the networks not shown in the figure but on the other side of the routers? Give reasons for your answers.

Scenario Answers

The answers provided in this section are not necessarily the only possible answers to the questions. The questions are designed to test your knowledge and to give practical exercise in certain key areas. This section is intended to test and exercise skills and concepts detailed in the body of this chapter.

If your answer is different, ask yourself whether it follows the tenets explained in the answers provided. Your answer is correct not if it matches the solution provided in the book but rather if it has included the principles of design laid out in the chapter.

In this way, the testing provided in these scenarios is deeper: It examines not only your knowledge but also your understanding and ability to apply that knowledge to problems.

If you do not get the correct answer, refer back to the text and review the subject tested. Be certain to also review your notes on the question to ensure that you understand the principles of the subject.

Scenario 14-1 Answers

1. *What needs to be done in addition to solve the problems caused by EIGRP? Give the configuration commands necessary to activate this solution on Router A.*

The WAN is a Frame Relay cloud, and Router A is the hub in the hub-and-spoke configuration. Each VC is 56 kbps.

The other solution that the consultant suggested was to perform summarization to limit the query range of the routers. This would prevent the routes in the topology table being SIA, which seriously affects the performance of the network.

The commands required are as follows:

```
RouterA(config)# router eigrp 63
RouterA(config)# no auto-summary
RouterA(config)# network 10.0.0.0
RouterA(config)# int s0
RouterA(config-if)# ip summary-address eigrp 63 10.1.0.0 255.255.0.0
```

2. *Give the commands to configure Router A for EIGRP over this NBMA cloud.*

The configuration on Router A is as follows:

```
RouterA(config)# interface serial 0
RouterA(config-if)# frame-relay encapsulation
RouterA(config-if)# bandwidth 178
```

3. *Give the commands to configure Router B for EIGRP over this NBMA cloud.*

The configuration on Router B is as follows:

```
RouterB(config)# interface serial 0
RouterB(config-if)# frame-relay encapsulation
RouterB(config-if)# bandwidth 56
```

Scenario 14-2 Answers

1. *How much bandwidth has each circuit been allocated? Why was this value chosen by the administrator?*

Because a maximum of 256 kbps is available, you cannot allow any individual PVC to handle more than 25 kbps (256/10). Note that EIGRP will not use more than 22.5 kbps (90 percent of 25 kbps) on this interface, even though its capacity is 56 kbps. This configuration will not affect user data capacity, which will still be able to use the entire 56 kbps, except when EIGRP is using 22 kbps.

2. *What bandwidth utilization is available to EIGRP? Why was this value chosen by the administrator?*

Because this data rate is low, and because you do not expect very much user data traffic, you can allow EIGRP to use up to 90 percent of the bandwidth.

3. *If Router A fails, what would the effect be on the network?*

If Router A fails, there would be no communication between the routers in the WAN because Router A is the hub. Each site would function, but they would all be isolated from each other. The neighbor tables would fail to hear the Hellos from the other routers connecting to the WAN and would time out all routes that they had heard from these routers. The topology table would be updated, and the routers would send updates to all their other neighbors.

4. *Is summarization possible only on the routers entering the WAN cloud, or is it possible on the networks not shown in the figure, but on the other side of the routers? Give reasons for your answers.*

Summarization is possible on all interfaces in EIGRP, as long as the addressing scheme allows for it to be implemented. This is one of the major advantages of EIGRP over OSPF. OSPF can summarize only at ABRs.

Part V: BGP

Chapter 15 Connecting to Other Autonomous Systems—The Basics of BGP

Chapter 16 Implementing and Tuning BGP for Use in Large Networks

Part V covers the following Cisco BSCI exam topics:

- Describe the features and operation of BGP
- Explain how BGP policy-based routing functions within an autonomous system
- Describe the scalability problems associated with internal BGP
- Given a set of network requirements, identify the steps to configure a BGP environment and verify proper operation (within described guidelines) of your routers
- Interpret the output of various **show** and **debug** commands to determine the cause of route selection errors and configuration problems



This chapter covers the following topics, which you need to understand to pass the CCNP/CCDP/CCIP BSCI exam:

- Introduction to BGP and communicating with other autonomous systems
- Overview of the BGP operation
- Basic configuration commands to connect to another autonomous system
- Managing and verifying the BGP configuration
- Alternative methods of connecting to another autonomous system

Connecting to Other Autonomous Systems— The Basics of BGP

This chapter details Border Gateway Protocol version 4 (BGP-4, referred to throughout this book as BGP). The CCNP/CCDP/CCIP BSCI exam barely scratches the surface of the detail available to this protocol. This chapter deals with the basic concepts and configuration commands of BGP, which are covered in greater complexity in the next chapter. This chapter builds on your understanding of routing within large enterprise networks. The introduction to BGP in this chapter will cover external BGP (eBGP) and familiarize you with the necessary terms and concepts.

This chapter is broken into two major areas. The first part deals with how BGP works theoretically. Implementing and managing a BGP network is described at the end of the chapter.

“Do I Know This Already?” Quiz

The purpose of the “Do I Know This Already?” quiz is to help you decide what parts of this chapter to use. If you already intend to read the entire chapter, you do not necessarily need to answer these questions now.

The 14-question quiz, derived from the major sections in the “Foundation Topics” portion of the chapter, helps you to determine how to spend your limited study time.

Table 15-1 outlines the major topics discussed in this chapter and the “Do I Know This Already?” quiz questions that correspond to those topics.

Table 15-1 “Do I Know This Already?” Foundation Topics Section-to-Question Mapping

Foundation Topics Section	Questions Covered in This Section
Introduction to BGP and Communicating with other Autonomous Systems	1–3
Overview of the BGP Operation	4–6
Basic Configuration Commands to Connect to Another Autonomous System	7–9
Managing and Verifying the BGP Configuration	10–12
Alternative Methods of Connecting to an Autonomous System	13–14

CAUTION The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark this question wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

1. In BGP, what is an attribute?
 - a. Topology table criteria
 - b. Characteristics of a path, similar to a metric
 - c. Links in an update
 - d. Neighbor table entry
2. Which of the following are key features of BGP?
 - a. Periodic keepalives
 - b. Periodic updates
 - c. Poison reverse
 - d. Triggered updates
3. What is the routing protocol eBGP used to communicate between?
 - a. Routers within an autonomous system
 - b. Routers in different countries
 - c. Routers in different autonomous systems
 - d. Routers running different IGPs
4. What is the transport protocol for BGP?
 - a. BGP runs directly on top of the data-link layer
 - b. IP
 - c. UDP
 - d. TCP
5. When are full routing updates sent in BGP?
 - a. At the beginning of each session
 - b. Whenever a fault is seen on the link
 - c. Every 30 minutes
 - d. At the startup of the BGP process

6. Which of the following are BGP message types?
 - a. Updates
 - b. Keepalives
 - c. ACK
 - d. Notification

7. What is the purpose of the **network** command?
 - a. To advertise the stated network
 - b. To identify the interfaces to run BGP
 - c. To forward stated networks
 - d. To create neighbors with within these networks

8. What is the purpose of the **neighbor** command?
 - a. Forces the path to be chosen
 - b. Identifies the next hop router and sends LSAs and Hellos
 - c. States the address and autonomous system of the neighbor with whom to peer
 - d. States the neighbor that might be considered as feasible paths

9. What command enables the BGP process?
 - a. **router bgp remote** *autonomous-system-number*
 - b. **router bgp** *process-id*
 - c. **ip routing bgp**
 - d. **router bgp** *autonomous-system-number*

10. Which command is used to show the BGP connections between peers?
 - a. **show ip bgp connections**
 - b. **show ip bgp neighbor**
 - c. **show ip bgp sessions**
 - d. **show ip bgp topology**

11. Which command is used to show all BGP connections?
 - a. **show ip bgp connections**
 - b. **show ip bgp sessions**

- c. **show ip bgp topology**
 - d. **show ip bgp summary**
12. What is the purpose of the command **clear ip bgp** ?
- a. To disconnect all sessions
 - b. To clear the BGP routing table
 - c. To end an administrative session
 - d. To clear all IGP entries from the routing table
13. Which of the following are valid alternatives to using BGP when connecting to another autonomous system?
- a. Using FTP
 - b. Default routes
 - c. Static routes
 - d. Merging the two autonomous systems into one
14. What is the administrative distance of a static route that has been configured to point to the outgoing interface?
- a. 0
 - b. 100
 - c. 1
 - d. 255

The answers to the “Do I Know This Already?” quiz are found in Appendix A, “Answers to Chapter ‘Do I Know This Already?’ Quizzes and Q&A Sections.” The suggested choices for your next step are as follows:

- **7 or less overall score** —Read the entire chapter. This includes the “Foundation Topics” and “Foundation Summary” sections, the “Q&A” section, and the “Scenarios” at the end of the chapter.
- **8–10 overall score** —Begin with the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. If you have trouble with these exercises, read the appropriate sections in “Foundation Topics.”
- **11 or more overall score** —If you want more review on these topics, skip to the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. Otherwise, move to the next chapter.

Foundation Topics

Introduction to BGP and Communicating with Other Autonomous Systems

BGP is an extremely complex protocol used throughout the Internet and within multinational organizations. Its main purpose is to connect very large networks or autonomous systems. Large companies can use BGP as the glue between countries in which they operate; for example, a government might use BGP among the divisions of its administration, and the military might use it among the Army, the Navy, and the Air Force.

The goal of an exterior routing protocol such as BGP is not to find a specific network, but to provide you with information that allows you to find the autonomous system wherein the network lies. The interior routing protocol, such as EIGRP, RIP, OSPF, and IS-IS, will find the specific network that you seek.

This section presents the basics of BGP, introduces the terminology used in BGP, and explains the role of BGP as a routing protocol in the IP community. It is important to understand when to use BGP and when to use other solutions, such as static and default routes. Within this introductory section, the main characteristics of BGP are explained, laying the foundation for the more complex explanation in the rest of the chapter.

Characteristics of BGP

The main characteristics of BGP can be distilled into a few short points. These characteristics demonstrate why this protocol is best suited to exterior routing. The interaction of these components and their relative importance to the operation or design of a BGP network is discussed in the rest of the chapter.

The key features of BGP include:

- It is a path vector routing protocol.
- BGP supports variable-length subnet mask (VLSM), classless interdomain routing (CIDR), and summarization.
- Full routing updates are sent at the start of the session; triggered updates are sent subsequently.
- It creates and maintains connections between peers, using TCP port 179.
- The connection is maintained by periodic keepalives.

- The failure to see a keepalive, an update, or receipt of a notification is the means by which destination networks and paths to those destinations are tracked. Any change in the network results in a triggered update.
- The metrics used in BGP are intricate and are the source of its complexity and its strength. The metric, referred to as *attributes*, allows great granularity in path selection.
- The use of hierarchical addressing and the capability to manipulate traffic flow results in a network that is designed to grow.
- It has its own routing table, although it is capable of both sharing and inquiring about the interior IP routing table.
- It is possible to manipulate the traffic flow by using attributes. Despite the complexity offered in path selection using policy-based routing, the traffic is still forwarded using the hop-by-hop paradigm. This means that no router can send traffic on a route that the next-hop router would not choose for itself.

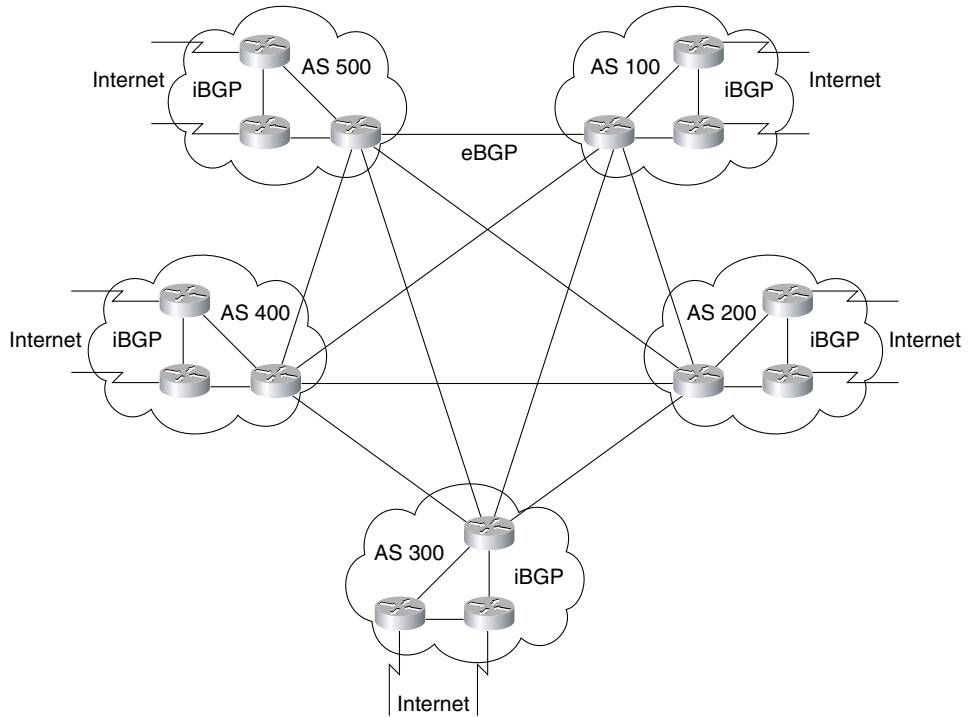
One of the most distinctive characteristics of BGP is its routing updates. When you look at the BGP updates, you can see why BGP is excellent for communicating between autonomous systems. BGP works as an exterior routing protocol because the routing updates are extremely concise. BGP is not interested in communicating a full knowledge of every subnet within the organization, but only in conveying enough information to find another autonomous system. The BGP routing update takes summarization to the extreme by communicating only a list of autonomous system numbers, aggregated prefix addresses, and some policy-based routing information.

However, the small amount of information that the BGP update carries is extremely important. BGP ensures the reliability of the transport carrying the updates and that the databases are synchronized.

BGP can be implemented in several ways, including between autonomous systems or across an autonomous system. When you use BGP to connect autonomous systems, it acts as an Exterior Gateway Protocol (EGP) and is referred to as External BGP, or eBGP. The routing protocol can also carry this external information between eBGP speakers that are resident in the same autonomous system. This is called Interior BGP (iBGP).

Figure 15-1 illustrates how BGP can be used to connect autonomous systems.

Many terms are used in BGP. You need to understand the naming of the parts in order to comprehend the working of the protocol as a whole.

Figure 15-1 *Using BGP to Connect Autonomous Systems*

BGP Terminology

BGP is a very different routing protocol from the interior routing protocols that you have studied so far. This protocol is designed to connect autonomous systems, not connect subnets within an autonomous system. BGP, therefore, comes with its own terms and concepts. As with other routing protocols, understanding the jargon of the technology is halfway to understanding the protocol itself.

Table 15-2 provides definitions for keywords and terms introduced in this chapter. A complete listing, including all of the chapter terms and additional terms, can be found in the Glossary toward the end of this book.

Table 15-2 *BGP Terms*

Term	Definition
aggregation	The BGP term for summarization.
attribute	Similar to a metric, these are variables that describe characteristics of the path to the destination. Once defined, the characteristics can be used to make decisions as to which path should be taken, based on the characteristic or attribute of the paths available.
autonomous system	Definition for the organizational boundary. Routers share the same routing protocol within the organization. The autonomous system defines all the routers within an administrative domain, where each router has full knowledge of the subnets within the domain. If you are connecting directly to the Internet using BGP, the autonomous system number must be unique and obtained from the Internet addressing committees.
Exterior Gateway Protocol (EGP)	The generic term for a protocol that runs between autonomous systems. There is also a specific EGP protocol with this name, which was the precursor to BGP.
External BGP (eBGP)	eBGP sends routing information between different autonomous system.
Interior Gateway Protocol (IGP)	This is a routing protocol that runs within an autonomous system. In the past, the term <i>gateway</i> was used to define a router.
Internal BGP (iBGP)	This is when BGP is used within an autonomous system. The routers are not required to be physical neighbors and often sit on edges of the autonomous system. iBGP is used between eBGP speakers in the same autonomous system.
Originator-ID	This is a BGP attribute. It is an optional nontransitive attribute that is created by the route reflector. The attribute contains the router ID of the router that originated the route in the update. The purpose of this attribute is to prevent a routing loop. If the originating router receives its own update, it ignores it.
policy-based routing	This allows the administrator to program the routing protocol by defining how traffic is routed. This is a form of static routing enforced by access lists called <i>route maps</i> . Policy-based routing (PBR) is protocol independent and uses route maps by creating a separate process to force routing decisions in a sophisticated way by matching and changing attributes and other criteria.
prefix list	The prefix list is used as an alternative to distribute lists to control how BGP learns or advertises updates. Prefix lists are faster, more flexible, and less processor-intensive than distribute lists.

Table 15-2 *BGP Terms (Continued)*

Term	Definition
route reflector	This is the router that is configured to forward routes from other identified iBGP clients. This removes the necessity for a fully meshed iBGP network, preserving network resources. A fully meshed network has a great deal of overhead and does not scale.
route reflector client	A client is a router that has a TCP session with its iBGP peer that is acting as a route reflector. It forwards routes to the route reflector, which propagates these on to other routers. The client does not have peer connections with other clients.
route reflector cluster	A cluster is a group consisting of a route reflector and clients. There can be more than one route reflector in a cluster.
synchronization rule	This rule states that a router cannot forward a route to an eBGP peer unless the route is in its local IP routing table. This requires that the IGP and BGP routing table are synchronized. This is to prevent BGP from advertising routes that the autonomous system cannot direct to the destination. If BGP is fully meshed, it does not have to rely on the IGP routing tables and you can disable synchronization. Use the router(config-router)#no synchronization command to turn off synchronization.
transit autonomous system	An autonomous system that is used to carry BGP traffic across to another autonomous system. None of the traffic is destined for any router within the transit autonomous system; it is simply being routed through it.

When to Use BGP

BGP is particularly complex when determining the path that should be taken or when used in conjunction with route maps to implement policy-based routing. Its complexity is its very strength. BGP is not a protocol for all occasions; its brevity and path determination make it a specialized routing protocol. There are specific situations where using BGP is important, including the following:

- Your organization is connecting to multiple ISPs or autonomous systems and is actively using those links. Many organizations use multiple links for redundancy purposes, justifying the additional cost by using all the links and reducing bottlenecks and congestion. In this case, policy-based routing decisions might need to be made on a link-by-link basis.
- The routing policy of the ISP and your company differ. The cost of the link depends on usage in addition to other factors that might need to be programmed into the BGP configuration to make the best use of the connection by manipulating the traffic.

- The traffic in your organization needs to be distinguished from that of the ISP. The two organizations cannot logically appear as one autonomous system.
- Your organization is an ISP and, therefore, conforms to criteria in the preceding bullets. The nature of your business requires the traffic from other autonomous systems to travel across your autonomous system, treating it as a transit domain.

When Not to Use BGP

A simple network is a network that is easier to manage and maintain, which is the main reason to avoid BGP configuration in the network. Therefore, if your network has the following characteristics, use other methods, such as static and default routing, to achieve connectivity to the ISP or to another autonomous system network:

- The ISP and your organization have the same routing policy.
- Although your company has multiple links to the ISP, these links are redundant and there are no plans to activate more than one link to the Internet.
- There are limited network resources, such as memory and CPU, on the routers.
- The bandwidth between the autonomous systems is low, and the additional routing overhead would detract from routing data.

How BGP works and an explanation of some of its characteristics are given in the following section.

Overview of the BGP Operation

BGP is connection-oriented. When a neighbor is seen, a TCP peering session is established and maintained. BGP probes are sent out periodically to sustain the link and maintain the session. These probes, otherwise known as *keepalives*, are simply the 19-byte header used in the BGP updates.

Having established the session, the routing tables are exchanged and synchronized. The routers then send incremental updates only when changes occur. The update refers to a single path and the networks that can be reached via that path. Having corrected the routing table, the BGP process propagates the change to all neighbors, with a few exceptions, based on an algorithm to ensure a loop-free network.

The operation of BGP is very straightforward. Indeed, all the complexity of the protocol is delivered in only a few different message types.

Four different message types are used in BGP:

- **Open messages** —Used to establish connections with peers.

- **Keepalives** — Sent periodically between peers to maintain connections and verify paths held by the router sending the keepalive. These packets are sent unreliably. If the periodic timer is set to a value of 0, this equates to infinity, and no keepalives are sent.
- **Update messages** — Contain paths to destination networks and the path attributes. Updates include routes that are no longer available or are withdrawn. There is one path per update, requiring many updates for many paths. The information contained in the update includes the path attributes such as origin, autonomous system path, neighbor, and interautonomous system metric.
- **Notification** — Used to inform the receiving router of errors that cause the connection to be closed.

The establishment of a BGP peer is shown in Figure 15-2.

The following sections discuss in greater detail how BGP summarizes routes, manipulates the path selection process using policy-based routing, and selects a route.

CIDR and Route Aggregation

BGP needs to communicate some information between the autonomous systems, but perhaps not all of the information is needed. If the network is designed to support summarization, the amount of network resources required could be reduced, because fewer memory, bandwidth, and CPU resources would be consumed when updating the routing tables. In fact, by summarizing the IANA classful addresses, the Internet becomes a quicker and more reliable environment.

CIDR is one of the main solutions implemented in recent years. This is a method of consolidating addresses into a few summary addresses. Instead of an address having a mask to identify the network portion of the address, it has a prefix mask, which is simply a number that indicates the number of bits that have been allocated to the network. Summarization reduces the number of prefix bits and thus incorporates other networks that share those prefix bits into one address. Prefix masks are now used with subnets as well as IANA classful addresses. If an autonomous system has been allocated eight consecutive Class C addresses, it can be advertised as one network, which is an example of summarization at its best.

BGP will propagate the prefix and the prefix mask together, allowing not only for the design of a truly hierarchical network, but also for the streamlining of the network resources. A router can pass on the aggregated routes, although it is capable of aggregating routes itself. Therefore, a router can send aggregated routes, routes that have not been summarized, or a mixture of both.

The process of how BGP aggregates routes in compliance with CIDR is shown in Figure 15-3. In the figure, the update sends 1 network, instead of 16 networks, to the router in autonomous system 200.

Figure 15-2 Establishing a BGP Peer

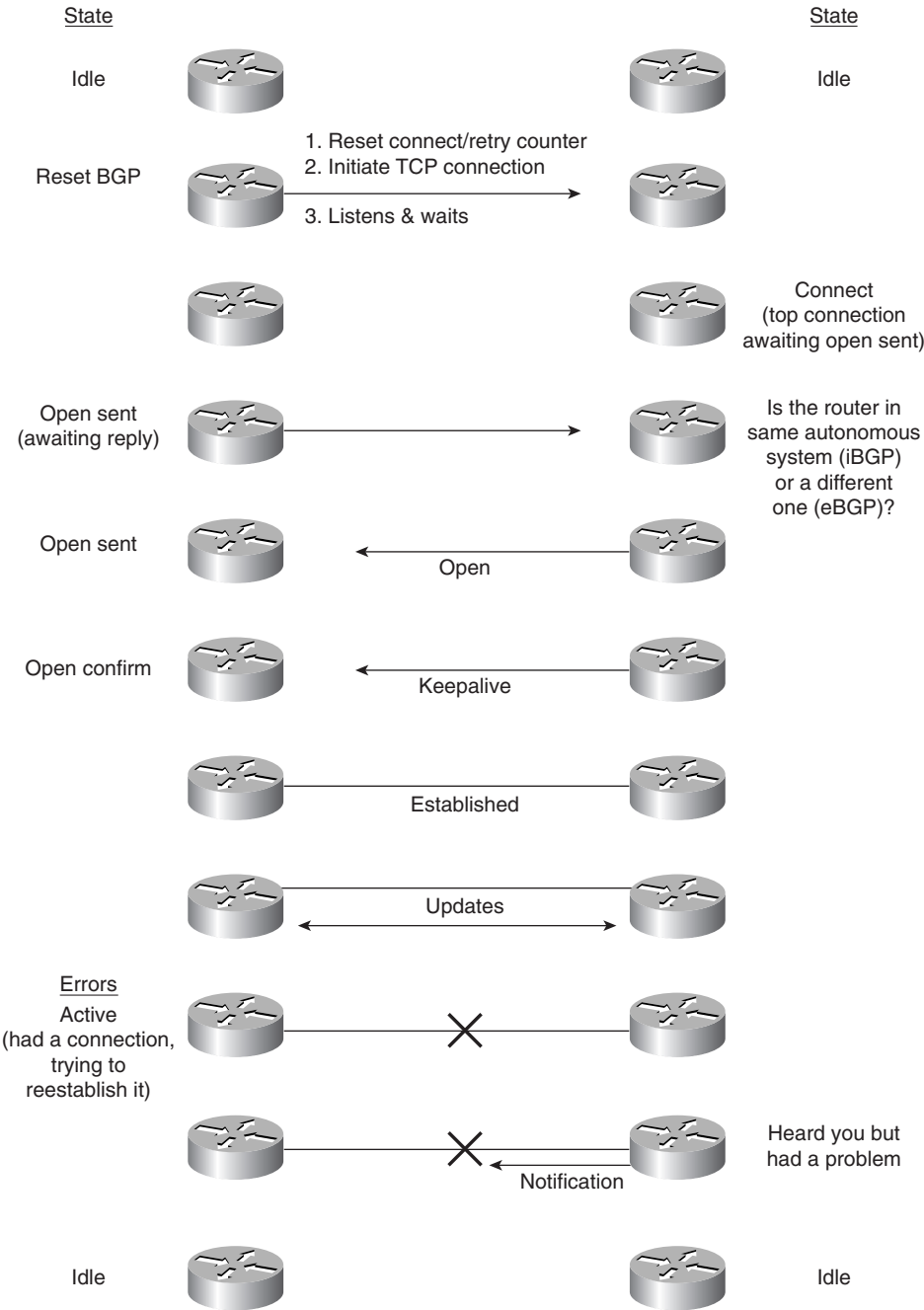
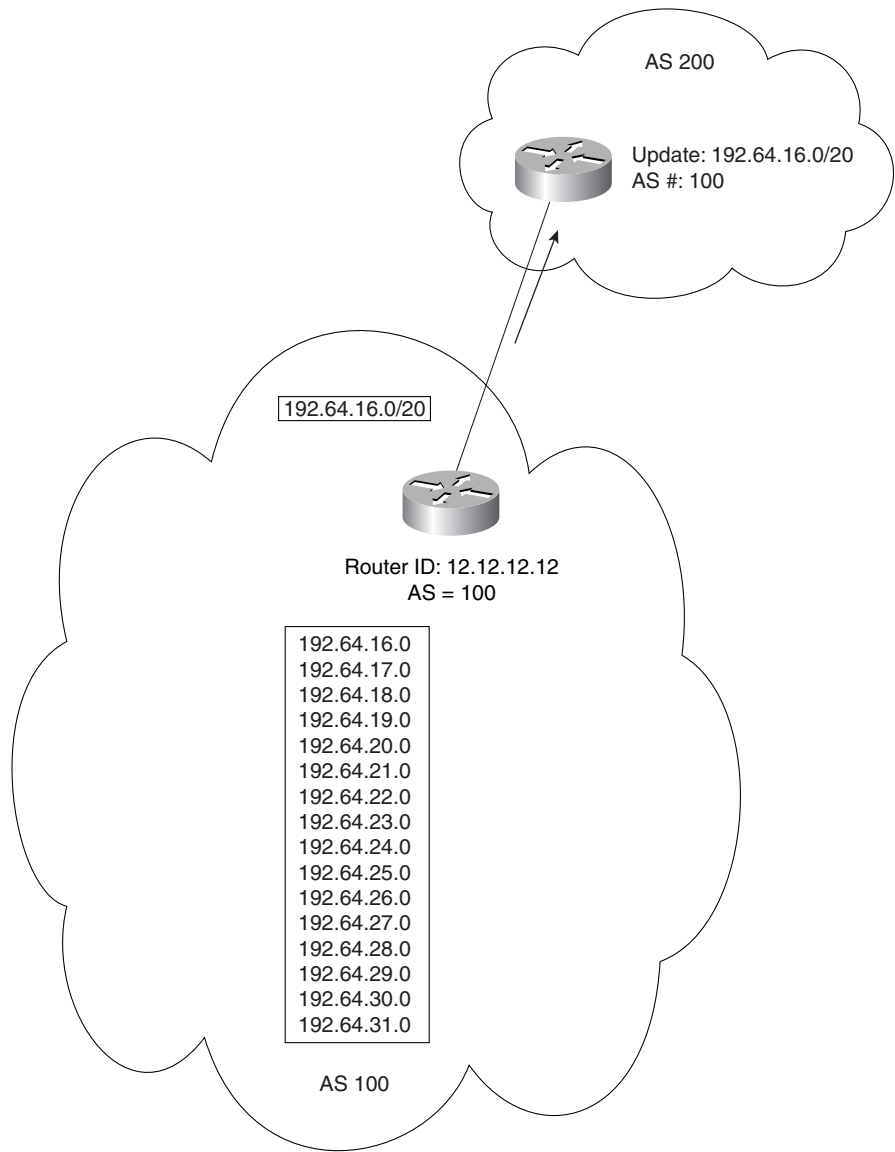


Figure 15-3 BGP and CIDR



BGP and Policy-Based Routing

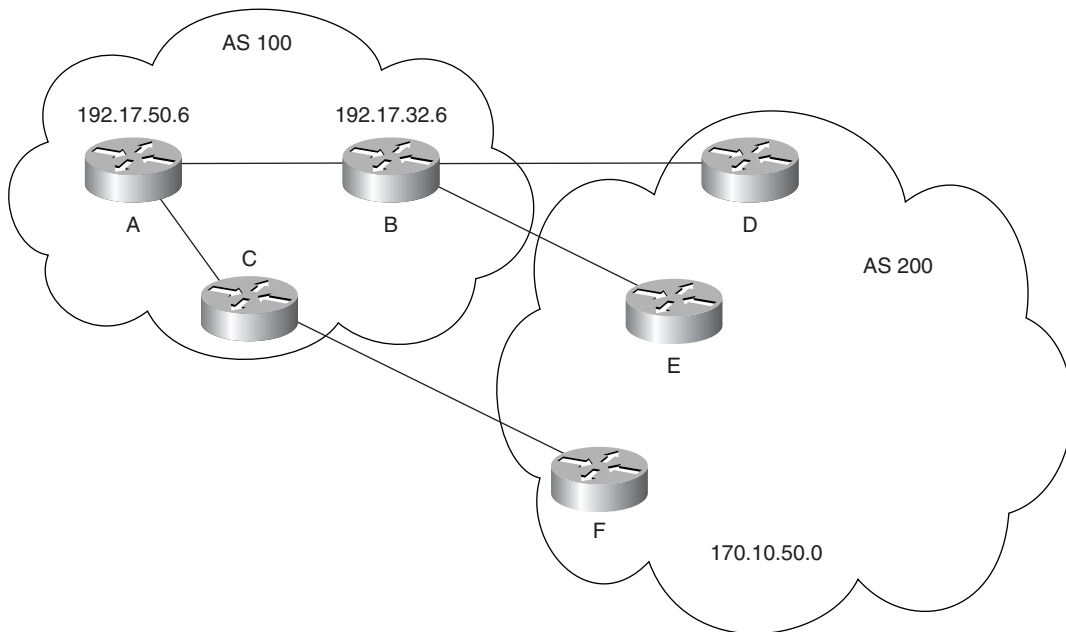
Policy-based routing is a separate administrative function from BGP, and they do not rely upon each other to work. Policy-based routing gives the administrator the ability to define how traffic will be routed at the autonomous system level. Policy-based routing is a form of static routing enforced by access lists, including route maps, distribute lists, prefix lists, and filter lists, because they impact

the inbound or outbound routing table updates. This is a level of control above the dynamic routing protocol. Given that many variables, or attributes, in BGP can influence dynamic routing, the administrator is provided with a very high level of control. This dimension distinguishes BGP from other routing protocols. You will learn more about BGP attributes later in this section.

One of the most sophisticated forms of policy-based routing is implemented with route maps, which are the topic of Chapter 18, “Controlling Network Traffic with Route Maps and Policy-Based Routing.” BGP’s capability to choose the routing path via conditional programming was used for policy routing in other situations before it became an option in the Cisco IOS software. BGP was deployed in some situations to use the policy routing-based options; for example, BGP was used by the two early deployments of tag routing as the only means of programming policy-based routing. This is a powerful tool and can be used in many situations, such as forcing traffic entering your routing domain to pass through a firewall or load balancing among multiple connections to the Internet. Refer to Figure 15-4 for an example of how policy-based routing could be implemented.

In Router A in Figure 15-4, the traffic from 192.17.50.6 is from the graphic design department. It is high-volume, sensitive traffic. Therefore, you should send it on a path dedicated to such traffic, so this traffic should be directed to Router C.

Figure 15-4 *Example of Policy-Based Routing Using BGP*



Policy-based routing is discussed in the following sections. You will be introduced to the rules that govern policy-based routing, the disadvantages that you should be aware of before implementing

this technology, and the concept of attributes and how attributes are used in policy-based routing to select a path across a nonbroadcast cloud.

Rules of Policy-Based Routing

BGP can implement any of the following rules associated with the hop-by-hop paradigm. This paradigm is the capability to influence which router will be the next-hop router, potentially dictating it at every router and thus influencing the entire path of the traffic, hop by hop. The following rules associated with policy-based routing seem repetitive, but, in fact, each point raises a subtly different nuance:

- Traffic can be directed on either the source address or both the source and destination addresses. These are only some of the criteria that can be used.
- Policy-based routing affects only the next hop in the path to the destination.
- Policy-based routing does not affect the destination of the packet. It affects the path used to get to the destination.
- Policy-based routing does not allow traffic sent into another autonomous system to take a different path from the one that would have been chosen by that autonomous system.
- It is possible to influence only how traffic will get to a neighboring autonomous system, not how it will be routed within that autonomous system.
- As policy-based routing examines the source address, it is configured on the inbound interface.

NOTE Some things should be considered before arbitrarily deciding to implement policy-based routing. Understand that any additional configuration requires additional CPU, particularly when every packet characteristic must be examined. It is also wise to have a backup path in place in case the defined next-hop router goes down. If there is no alternative defined, policy-based routing will default to dynamic routing decisions. Of course, whenever extra configuration is required, there is the possibility of misconfiguration or of disrupting other traffic. Refer to Chapter 18 for more information about the disadvantages of policy-based routing.

BGP Attributes

The key to BGP is the capability to divert traffic into different directions based on criteria determined by the network architects. BGP is concerned with providing the capability to manipulate the traffic flow through the network. This is the key difference between the routing protocols you have seen in previous chapters and BGP.

BGP directs the traffic flow by using attributes. This is different from policy-based routing, which is a sophisticated method of forcing traffic down a particular path in spite of the dynamic routing decisions. The use of attributes refers to the use of variables in the selection of the best path for the dynamic routing protocol BGP.

BGP uses attributes to select the best path. In essence, attributes are the metric for BGP. However, they are more than a list of variables by which a route is selected. They also carry information on which decisions are based, hence the name *attributes*. The variables describe characteristics or attributes of the path to the destination. These characteristics can be used to distinguish the paths, which is policy-based routing. So, policy-based routing uses the attributes in BGP to make sophisticated decisions in path selection.

The reason for choosing different paths is determined by the nature of the traffic. Policy-based routing allows you to select reliable and slow links for packets that need to be propagated.

Some of the information carried in the update messages is more important than others. Indeed, some of this information is crucial to the successful operation of BGP, so it must be carried through the network to every router running BGP. The path selection is based on the attributes and their values.

Because the BGP information in the updates varies in significance to the BGP network, it has been categorized by importance. The attributes are divided into two types, well-known and optional. The well-known attributes are those attributes whose implementation is mandatory, whereas the optional ones are just that—optional. Both of these are subdivided into two further categories, allowing considerable granularity (see Table 15-3).

Table 15-3 The Four Categories of Attributes

Category	Description
Well-known:	
Mandatory (required by all routers)	These attributes are required and are therefore recognized by all BGP implementations.
Discretionary (required by all routers and recognized by all routers)	It is not required that these attributes be present in the update messages, but if they are present, all routers running BGP will recognize and act on the information contained.
Optional:	
Transitive	The router might not recognize these attributes, but if this is the case, it marks the update as partial and sends the update, complete with attributes, to the next router. The attributes traverse the router unchanged, if they are not recognized.
Nontransitive	Nontransitive attributes are dropped if they fall onto a router that does not understand or recognize the attribute. These attributes will <i>not</i> be propagated to the BGP peers. Unrecognized nontransitive optional attributes must be quietly ignored and not passed along to other BGP peers. New transitive optional attributes might be attached to the path by the originator or by any other autonomous system in the path (see RFC 1771).

The attributes are appropriately carried in the updates that inform the routers of the routes.

Although there are other attributes, the following list includes the ones supported by Cisco. The attributes and a description of their characteristics are listed in Table 15-4 for quick reference and comparison.

Table 15-4 *The BGP Attributes*

Attribute Name	Category	Code	Preference	Description
Origin	Well-known, mandatory	1	Lowest origin code Where: IGP < EGP < Incomplete	This path attribute identifies the source of the routing update. The possible sources of routing information are as follows: • The path originates from within the autonomous system. It was created with the iBGP network command. The route will be marked in the BGP routing table with an “i.” • If the source is an exterior routing protocol, it will be identified with an “e” in the routing table. • The route could have been redistributed into BGP, so there is incomplete information. The route is marked by a question mark (?).
AS_Path	Well-known, mandatory	2	Shortest path	This attribute includes a list of all the autonomous systems that the routes have passed through.
Next Hop	Well-known, mandatory	3	Shortest path or IGP metric	The next hop attribute states the next hop on the path for the router to take. In eBGP, this is the source address of the router that sent the update from the other autonomous system. In iBGP, for routes originated outside the autonomous system, the address is still the source address of the router that sent the update. The protocol states that the next hop advertised by eBGP should be carried into iBGP. Therefore, it is important for the IGP to be aware of this network so that any router within the autonomous system can reach the next hop.

(continues)

Table 15-4 *The BGP Attributes (Continued)*

Attribute Name	Category	Code	Preference	Description
Multiple Exit Discriminator (MED)	Optional, nontransitive	4	Lowest value	This attribute informs routers outside the autonomous system which path to take into the autonomous system. The MED is known as the external metric of a route. Therefore, the MED is passed between the autonomous systems, but it will not be propagated into a third autonomous system.
Local preference	Well-known, discretionary	5	Highest value	This attribute is used to tell routers within the autonomous system how to exit the autonomous system in the case of multiple paths. It is the opposite of the MED attribute. This value is passed solely between iBGP peers.
Atomic aggregate	Well-known, discretionary	6	Information not used in path selection	The originator of the aggregate route is stated. It is useful because it shows that some information has been lost because of the aggregation of routes. It states to the receiving router in another autonomous system that the originator of the route aggregated the routes.
Aggregator	Optional, transitive	7	Information not used in path selection	This attribute states the BGP router ID and autonomous system number of the router that was responsible for aggregating the route. This attribute will include a list of all the autonomous systems that the aggregated routes have passed through.

Table 15-4 *The BGP Attributes (Continued)*

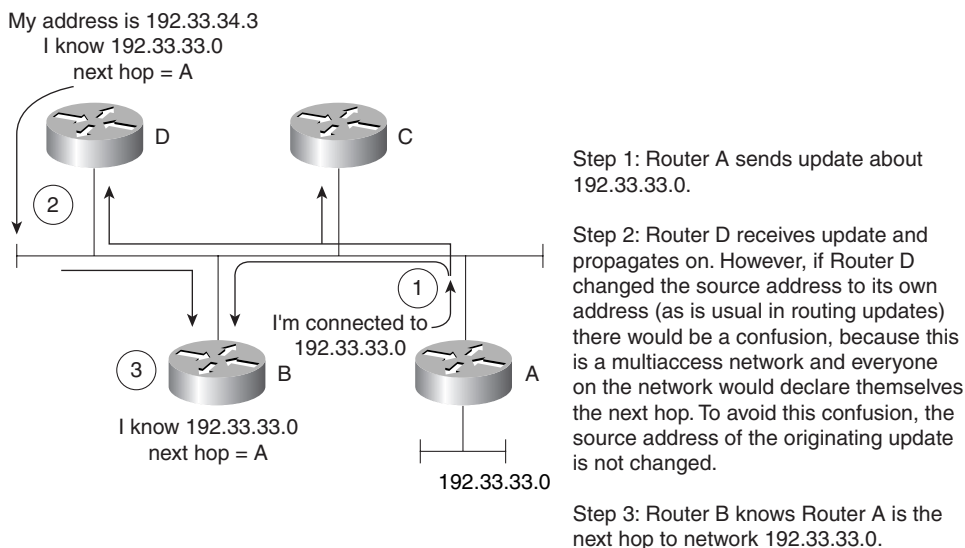
Attribute Name	Category	Code	Preference	Description
Community	Optional, transitive	8	Information not used in path selection	This is the capability to tag certain routes that have something in common. They are thereby made members of the same club or community. This is often used in conjunction with another attribute that will affect route selection for the community. For example, the use of the local preference and community attributes would allow the network administrators and other privileged employees to use the high-speed link to the Internet, while others share a fractional T1. Communities have no geographical or logical limits. BGP can filter on incoming or outgoing routes for filtering, redistribution, or path selection.
Originator ID	Optional, nontransitive	9	Information not used in path selection	The route reflector (described in Chapter 16, “Implementing and Tuning BGP for Use in Large Networks”) appends this attribute. It carries the router ID of the originating router in the local autonomous system. It is used to prevent loops.
Cluster ID	Optional, nontransitive	10	Information not used in path selection	The cluster identifies the routers involved in the route reflection. The cluster list shows the reflection path that has been taken. This is used to prevent looping errors.
Weight	Cisco-defined		Highest value	This is proprietary to Cisco and is used in route selection. It is local to the router, and because it is not propagated to other routers, there is no problem with compatibility. When there are multiple paths, it selects a path to a destination with different next hops to the same destination. Note that the weight attribute has no code. Because it is a local attribute and is not propagated to other routers, no code is needed.

The Next-Hop Attribute and a Broadcast Multiaccess Network

It is worth mentioning a potential problem with one of the attributes, namely the next-hop attribute. In eBGP, the next hop is the IP address of a router from outside the autonomous system advertising into the autonomous system. The IP address is that specified in the **neighbor** command. However, on a multiaccess network, if a route came from one router, it would be unwise for another router to readvertise the route with its own address as the source address. This might lead to packets being sent to several routers on the multiaccess network before finding the true originator, which is obviously inefficient. The rule, therefore, is that the address of the router that originally sent the update onto the multiaccess network should remain as the source address.

Figure 15-5 shows the next-hop router in a multiaccess network.

Figure 15-5 *The Next-Hop Router in a Multiaccess Network*



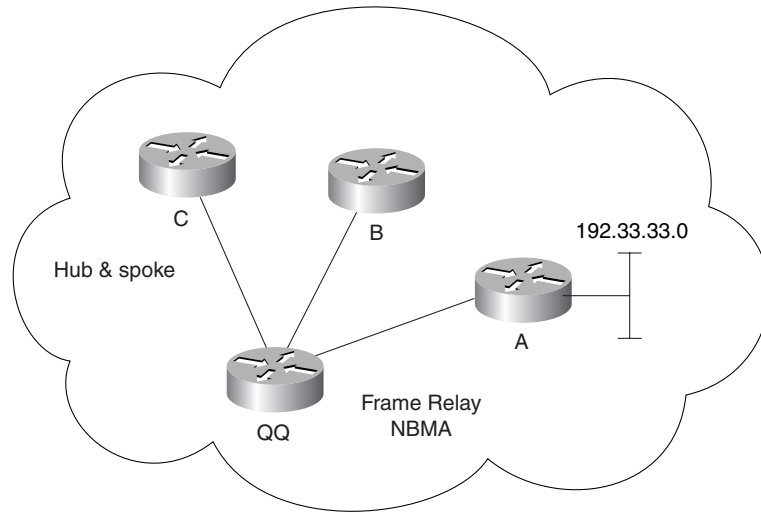
As shown in Figure 15-5, the source address is not changed, so the next-hop address is always the source (in this case, Router A). When this rule is applied to nonbroadcast multiaccess (NBMA), problems arise because although Router B might correctly point to Router A, Router B cannot see Router A in this NBMA cloud. Extra configuration is needed, as described in the next section.

The Next-Hop Attribute and a Nonbroadcast Multiaccess Network

As a multiaccess network, an NBMA network plays by the same rules of maintaining the source address of the router that originated the route on the network. However, there is a potential problem because the other routers are not going to be communicating directly with the source router if the NBMA cloud has a hub-and-spoke configuration (see Figure 15-6). If this is the case, the problem

can be solved with a command that forces the router to advertise itself as the source. The commands for configuring BGP are dealt with later in the chapter in the section “Defining the Networks to Be Advertised.”

Figure 15-6 *The Next-Hop Router in an NBMA Network*



Route Selection Process

As you have seen, BGP has many options by which to select one route over another using attributes. Now that you understand how attributes work in BGP, this section outlines the overall process of route selection.

The following process illustrates the logic of the selection process:

1. If the router has a valid route to the destination, use that route.
2. If there is more than one valid route to the destination, take the route with the highest weight (Cisco proprietary).
3. If the weights are the same, select the route with the highest local preference.
4. If the routes have the same local preference, prefer the route that originated on that router.
5. If there are no routes that originated on the router, examine the AS_Path and select the shortest path (through the fewest autonomous systems).
6. If the AS_Path is the same, examine and choose the lowest origin code, where an internal path has a lower value than an external path, which in turn is lower than an incomplete path.
7. If the origin codes are the same and the routes originate from a common autonomous system, select the path with the lowest MED.

8. If the MED values are the same, choose an external BGP route over an internal BGP route.
9. If there is no external route, choose the path with the lowest IGP metric or cost to the next-hop router for iBGP.
10. Is one of the routes older? If so, choose the most recent route.
11. If all else fails, choose the router with the lowest BGP router ID.

NOTE The preceding is a simplification of the selection process. For more detailed information on how the routing decisions are made, refer to RFC 1771, “A Border Gateway Protocol 4 (BGP).”

This document, together with its companion document, RFC 1772, “Application of the Border Gateway Protocol in the Internet,” defines an interautonomous system routing protocol for the Internet.

Basic Configuration Commands to Connect to Another Autonomous System

To connect to another autonomous system, it is necessary to configure required commands. You can also configure optional commands based on the needs of your network environment. The next sections discuss both required and optional commands.

Configuring Required BGP Commands

To connect to another autonomous system, you need to configure the following items:

- The start of the routing process
- The BGP neighbor with which the routing process will synchronize routing tables over a TCP session

Starting the Routing Process

The command to configure the routing process is the same command as that for the interior routing protocols. The syntax is as follows:

```
Router(config)#router bgp autonomous-system-number
```

Identifying Neighbors and Defining Peer Groups

A peer group is a group of neighbors that share the same update policy. This is the mechanism by which routers are grouped to simplify configuration.

A peer group also eases the overhead on the network, because the iBGP routers do not need to be fully meshed. When the network is fully meshed, the maintenance of all the neighbor relationships can place strain on your network.

All the members of the peer group will inherit the characteristic configured with the following command. The use of *peer-group-name* allows the identification of the router as a member of a peer group:

```
Router(config-router)#neighbor peer-group-name peer-group
```

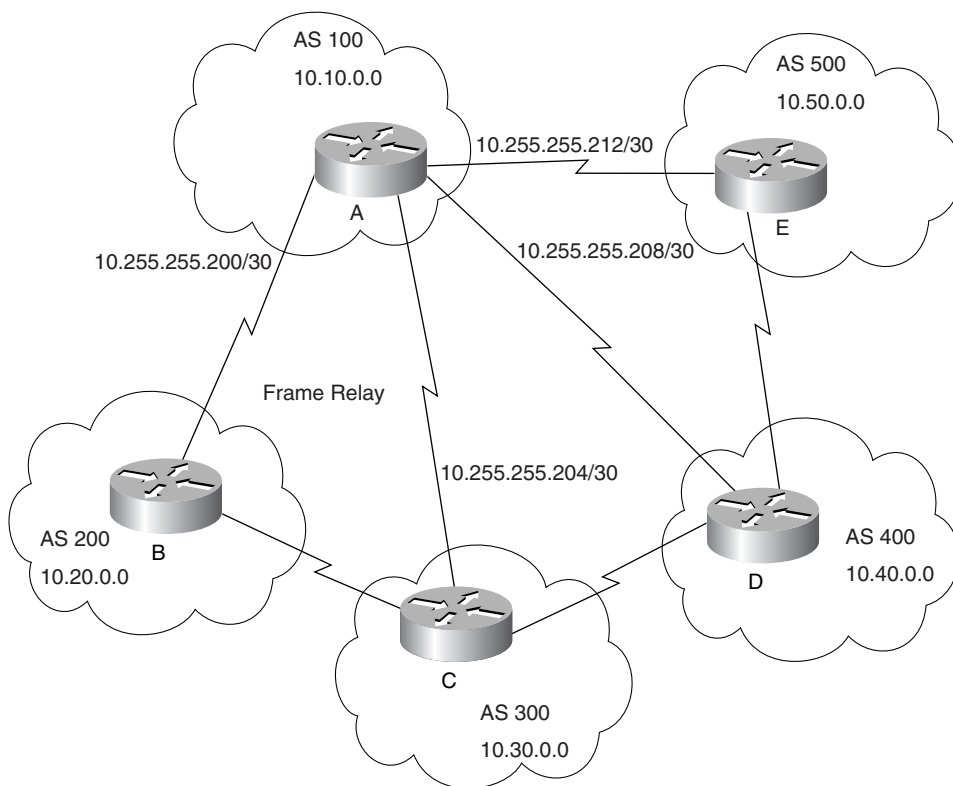
Once the peer group has been defined, it is possible to define neighbors for all the members of the peer group.

```
Router(config-router)#neighbor ip-address | peer-group-name remote-as  
autonomous-system-number
```

In iBGP, the remote autonomous system numbers that are defined for the BGP peers and for this BGP routing process are the same; in eBGP, these numbers differ.

Figure 15-7 is the topology map for Example 15-1. This figure shows the basic configuration commands required to make eBGP operate between autonomous systems.

Figure 15-7 Topology for Example 15-1



In Example 15-1, Router A in AS100 connects to routers in AS100, AS200, AS300, AS400, and AS500.

Example 15-1 *Basic Configuration of eBGP on Router A*

```
Router A

Router(config)#interface Serial0.1
Router(config-int)#ip address 10.255.255.201 255.255.255.252
!
Router(config)#interface Serial0.2
Router(config-int)#ip address 10.255.255.205 255.255.255.252
!
Router(config)#interface Serial0.3
Router(config-int)#ip address 10.255.255.209 255.255.255.252
!
Router(config)#interface Serial0.4
Router(config-int)#ip address 10.255.255.213 255.255.255.252
!
Router(config)#router bgp 100
Router(config-router)#neighbor 10.255.255.202 remote-as 200
Router(config-router)#neighbor 10.255.255.206 remote-as 300
Router(config-router)#neighbor 10.255.255.210 remote-as 400
Router(config-router)#neighbor 10.255.255.214 remote-as 500
```

Configuring Optional BGP Commands

The following sections describe optional commands in the basic configuration of BGP. These commands are related to the following tasks:

- Defining the networks to be advertised
- Forcing the next-hop address
- Aggregating routes

Defining the Networks to Be Advertised

To define the network that is to be advertised for this autonomous system, use the following command (each network requires a separate command):

```
Router(config-router)#network network-number mask network-mask
```

The **network** command determines the networks that are originated by this router. This is a different use of the **network** command than you are accustomed to configuring with EIGRP, OSPF, and RIP. This command is not identifying the interfaces upon which to run BGP; instead, it is stating the

networks that are to be advertised by BGP. The **network** command must include all the networks in the autonomous system to be advertised, not just those that are directly connected to the router. The mask portion is used because BGP can handle subnetting and supernetting. The mask used is a network mask.

Forcing the Next-Hop Address

On a multiaccess network, the rule is that the source address of a packet is that of the router that originated the packet onto the network. This can cause problems on an NBMA network that appears to be a multiaccess network, but that in reality might not have full connectivity to all the routers on the network. If the source address is the address of the initiating router, the other routers might not have a path to this next hop, and packets will be dropped. (This issue was described earlier in this chapter in the section “The Next-Hop Attribute and a Broadcast Multiaccess Network.”) To overcome this problem, the next-hop address can be configured to be that of the transmitting router. The syntax of the command is as follows:

```
Router(config-router)#neighbor {ip-address | peer-group} next-hop-self
```

Aggregating Routes

To summarize or aggregate routes within the BGP domain, use the following command from config-router mode:

```
Router(config-router)#aggregate-address ip-address mask [summary-only] [as-set]
```

If the parameter **summary-only** is used, then the specific routes are suppressed, and the summary route is the only one propagated. If the parameter **as-set** is used, then all the autonomous systems that have been traversed will be recorded in the update message. The **as-path** attributes of the prefixes that make up the aggregate are placed in the **as-set** attribute of the update.

Because it might be necessary to redistribute BGP into the IGP, care must be taken to avoid routing loops and not to overwhelm the routing tables. The administrative distance of BGP helps prevent this problem (see Figure 15-8).

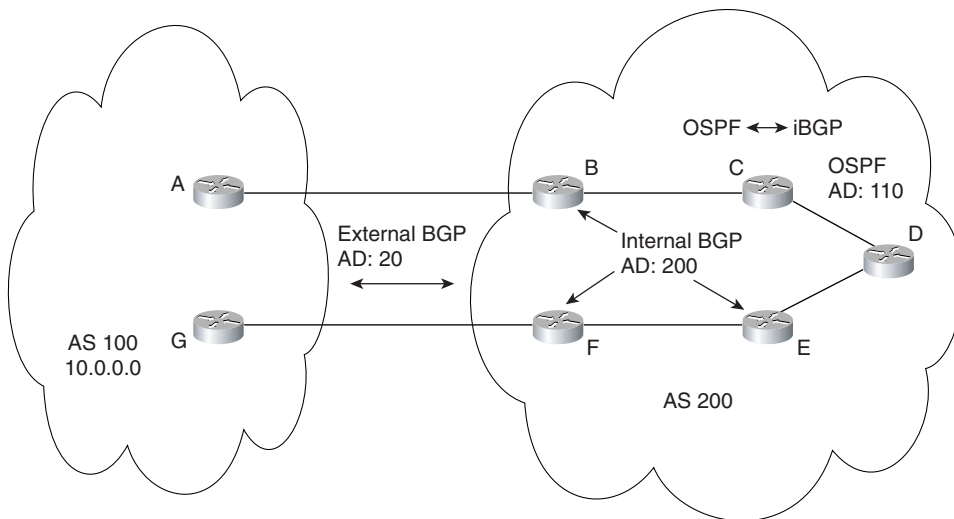
Example 15-2 is a BGP configuration using the **aggregate-address** command with the summarization parameter. In this example, any routes with a longer prefix than /16 are summarized, if possible, into the routes specified. In this way, the route 10.20.35.8/29 is not seen in BGP updates because it is summarized into the address 10.20.0.0/16. Any route that does not summarize into those specified in the example is listed individually. Summarization reduces the overhead on the network and simplifies system administration.

Example 15-2 *The aggregate-address Command with the Summarization Parameter*

```

Router A
Router(config)#interface Serial0
Router(config-int)#ip address 10.255.255.201 255.255.255.224
!
Router(config)#router bgp 100
!
Router(config-router)#network 10.10.0.0 255.255.0.0
!
aggregate-address 10.20.0.0 255.255.0.0 summary-only
aggregate-address 10.30.0.0 255.255.0.0 summary-only
aggregate-address 10.40.0.0 255.255.0.0 summary-only
aggregate-address 10.50.0.0 255.255.0.0 summary-only
!
Router(config-router)#neighbor 10.255.255.202 remote-as 200
Router(config-router)#neighbor 10.255.255.202 next-hop-self
!
Router(config-router)#neighbor 10.255.255.206 remote-as 300
Router(config-router)#neighbor 10.255.255.206 next-hop-self
!
Router(config-router)#neighbor 10.255.255.210 remote-as 400
Router(config-router)#neighbor 10.255.255.210 next-hop-self
!
Router(config-router)#neighbor 10.255.255.214 remote-as 500
Router(config-router)#neighbor 10.255.255.214 next-hop-self

```

Figure 15-8 *Administrative Distance and BGP*

In Figure 15-8, because iBGP has an administrative distance of 200, it is effectively the last choice. Router C will always choose the direct path to 10.0.0.0 through Router A because it has an administrative distance of 20. The route through D and B to find network 10.0.0.0 will have an administrative distance of 110 (OSPF) or 200 (iBGP), depending on the configuration.

Managing and Verifying the BGP Configuration

After configuration changes in BGP, it is necessary to reset the TCP session between neighbors. This can be forced with the following command:

```
Router(config-router)#clear ip bgp {* | address}[soft [in | out]]
```

This command disconnects the session between the neighbors and reestablishes it using the new configuration that has been entered. The **soft** option does not tear down the sessions, but it resends the updates. The **in** and **out** options allow the configuration of inbound or outbound soft updates. The default is for both.

NOTE The **clear ip bgp** command is an executive command executed at the privileged level. It is not a configuration command.

The **show** commands for BGP are comprehensive and give clear information about the BGP sessions and routing options. These informative commands and their functions are as follows:

- **show ip bgp** —Displays the BGP routing table.
- **show ip bgp paths** —Displays the topology table.
- **show ip bgp summary** —Displays information about the TCP sessions.
- **show ip bgp neighbors** —Displays information about the TCP connections to neighbors. When the connection is established, the neighbors can exchange updates.
- **show processes cpu** —Displays active processes and is useful to identify any process that is using excessive resources.

Another command that helps to troubleshoot any implementation and should be considered in the BGP configuration is the **debug** command. An entire book in the IOS documentation set is devoted to this command. For BGP, **debug** is a very useful command. It is shown here with all the possible options:

```
Router#debug ip bgp [dampening | events | keepalives | updates]
```

This command displays live information of events as they occur. The options available display dampening information, events, keepalives to maintain the TCP session with the peer, and routing updates as they are sent or received. *Route dampening* is a mechanism to minimize the instability caused by route flapping. By counting the number of times the identified route fails and applying a

penalty to the route for each flap, BGP categorizes the route as well-behaved or ill-behaved. Depending on its reputation, BGP suppresses the route by not advertising it in routing updates. The route can redeem itself by remaining up. Once there is some confidence that the route is stable, BGP reinstates and advertises it once more.

Given its complexity and role in internetworking, BGP is seldom used by private organizations. To handle the rush to connect to expanding Internet resources, service providers emerged to set up and manage the connection. This is advantageous for everyone. For a small fee, the organization or individual has a complex connection created and maintained. The burden on the Internet is also eased, because its limited space is taken by large corporations and organizations rather than millions of individual users. Because ISPs handle BGP, information on troubleshooting BGP is not covered in as much detail as has been done in previous chapters for other technologies.

Alternative Methods of Connecting to Another Autonomous System

If BGP is unnecessary in your network, consider the other possibilities for connecting to another autonomous system. If you do not have BGP running in your organization, you most likely are connecting to an ISP. This section is primarily concerned with connecting your company to the Internet through an ISP. The different methods include the following:

- A default route into the autonomous system and a static route from the autonomous system into the organization.
- A routing protocol into the autonomous system, making the autonomous system part of your autonomous system. The autonomous system will be using redistribution within its domain, and it is advisable for the organization to use some form of security, such as an access list or a firewall.

Typically, your ISP gives you a written sheet explaining the required configuration, or it requests access to your autonomous system boundary router (ASBR) so that it can configure the router itself. Either way, it is useful to be cognizant about the procedure.

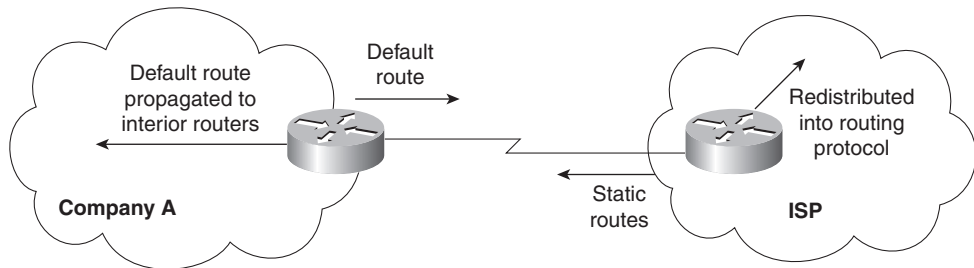
There are too many variations in configuring an internal routing protocol to detail in this chapter. Refer to Chapter 1, “IP Routing Principles,” for coverage of static and default routes; Chapter 3, “Designing IP Networks,” for connecting to the outside world with NAT; and Chapter 17, “Implementing Redistribution and Controlling Routing Updates,” for redistribution.

The use of default and static routes is an alternative to BGP that has been implemented for years in connecting remote satellite networks, particularly those connected by a dialup link.

The solution is simple: The smaller network defines a default route that it propagates throughout the domain. The default route points to the network that connects to the ISP. The larger autonomous

system—in this case, the ISP—configures static routes to its customer. The ISP's routing table summarizes the static routes to the IANA classful address and propagates the summarized routes throughout the ISP's network, eventually advertising them into the Internet. Thus, the smaller organization is connected into the global internetwork. Figure 15-9 illustrates the use of default and static routes and shows how they are propagated.

Figure 15-9 *Default and Static Route Configuration into the Internet*



The command syntax to configure a static route is as follows:

```
Router(config)#ip route prefix mask {interface | ip-address} [distance]
```

It is possible to specify either the outgoing interface or the IP address of the next logical hop. If you specify the outgoing interface, the routing table considers the route to be directly connected. The default administrative distance of a directly connected link is 0. This is only appropriate for point-to-point links where there is no choice in the destination address. If the next hop IP address is configured, the route is given the administrative distance of 1.

In either case, the administrative distance ensures that this path is chosen above all others. If this static route is a redundant link into the ISP configured as a backup resource, it might be necessary to change the administrative distance.

To propagate the static routes throughout the domain, it is necessary to redistribute them in to the appropriate routing protocol.

Foundation Summary

The “Foundation Summary” section of each chapter lists the most important facts from the chapter. Although this section does not list every fact from the chapter that will be on your exam, a well-prepared candidate should, at a minimum, know all the details in each “Foundation Summary” before going to take the exam.

The key features of BGP include the following:

- BGP is an enhanced path vector protocol.
- BGP supports VLSM, CIDR, and summarization.
- BGP creates and maintains connections between peers, using the destination TCP port 179 to initiate the connection.
- The connection is maintained by periodic keepalives.
- The failure to see a keepalive, an update, or the receipt of a notification is the means by which destination networks and paths to those destinations are tracked. Any change in the network results in a triggered update.
- The metric used in BGP is intricate and is the source of its complexity and its strength. The metric, referred to as *attributes*, allows great granularity in path selection.
- The use of hierarchical addressing and the capability to manipulate traffic flow results in a network that is designed to grow.
- BGP has its own routing table, although it is capable of both sharing and inquiring about the interior IP routing table.
- It is possible to manipulate the traffic flow by using the complex metric called *attributes*. Despite the complexity offered in path selection using policy-based routing, the traffic is still forwarded using the hop-by-hop paradigm. This means that no router can send traffic on a route that the next-hop router would not choose for itself.

Table 15-5 explains the four categories of attributes.

Table 15-5 *The Four Categories of Attributes*

Category	Description
Well-Known: Mandatory (required by all routers) Discretionary (required by all routers and recognized by all routers)	These attributes are required and are therefore recognized by all BGP implementations. These attributes are not required to be present in the update messages, but if they are present, all routers running BGP will recognize and act on the information contained.
Optional: Transitive Nontransitive	The router might not recognize these attributes, but if this is the case, it marks the update as partial and sends the update, complete with attributes, to the next router. The attributes traverse the router unchanged, if they are not recognized. Nontransitive attributes are dropped if they fall onto a router that does not understand or recognize the attribute. These attributes will <i>not</i> be propagated to the BGP peers. Unrecognized nontransitive optional attributes must be quietly ignored and not passed along to other BGP peers. New transitive optional attributes might be attached to the path by the originator or by any other autonomous system in the path (see RFC 1771).

The attributes are appropriately carried in the updates that inform the routers of the routes.

The attributes and a description of their characteristics are shown in Table 15-6.

Table 15-6 The BGP Attributes

Attribute Name	Category	Code	Preference	Description
Origin	Well-known, mandatory	1	Lowest origin code Where: IGP < EGP < Incomplete	This path attribute identifies the source of the routing update. The possible sources of routing information are as follows: • The path originates from within the autonomous system. It was created with the iBGP network command. The route will be marked in the BGP routing table with an “i.” • If the source is an exterior routing protocol, it will be identified with an “e” in the routing table. • The route could have been redistributed into BGP; as such, there is incomplete information. The route is marked by a question mark (?).
AS_Path	Well-known, mandatory	2	Shortest path	This is a sequence of the autonomous systems that the prefix has traversed.
Next hop	Well-known, Mandatory	3	Shortest path or IGP metric	The next hop attribute states the next hop on the path for the router to take. In eBGP, this will be the source address of the router that sent the update from the other autonomous system. In iBGP, for routes that originated outside the autonomous system, the address will still be the source address of the router that sent the update. The protocol states that the next hop advertised by eBGP should be carried into the iBGP. Therefore, it is important that the IGP is aware of this network so that any router within the autonomous system can reach the next hop.

Table 15-6 *The BGP Attributes (Continued)*

Attribute Name	Category	Code	Preference	Description
Multiple Exit Discriminator (MED)	Optional, nontransitive	4	Lowest value	This attribute informs routers outside the autonomous system which path to take into the autonomous system. It is known as the external metric of a route. Therefore, it is passed between the autonomous systems, but it will not be propagated into a third autonomous system.
Local preference	Well-known, discretionary	5	Highest value	This attribute is used to tell routers within the autonomous system how to exit the autonomous system in the case of multiple paths. It is the opposite of the MED attribute. This value is passed solely between iBGP peers.
Atomic aggregate	Well-known, discretionary	6	Information not used in path selection	This attribute states that the routes have been aggregated and that some information has been lost.
Aggregator	Optional, transitive	7	Information not used in path selection	This attribute states the BGP router ID and the autonomous system number of the router that was responsible for aggregating the route.
Community	Optional, transitive	8	Information not used in path selection	This is the capability to tag certain routes that have something in common. They are thereby made members of the same “club” or community. This is often used in conjunction with another attribute that will affect route selection for the community. For example, the use of the local preference and community attributes would allow the network administrators and other privileged beings to use the high-speed link to the Internet, while others shared a fractional T1. Communities have no geographical or logical limits. BGP can filter on incoming or outgoing routes for filtering, redistribution, or path selection.

(continues)

Table 15-6 The BGP Attributes (Continued)

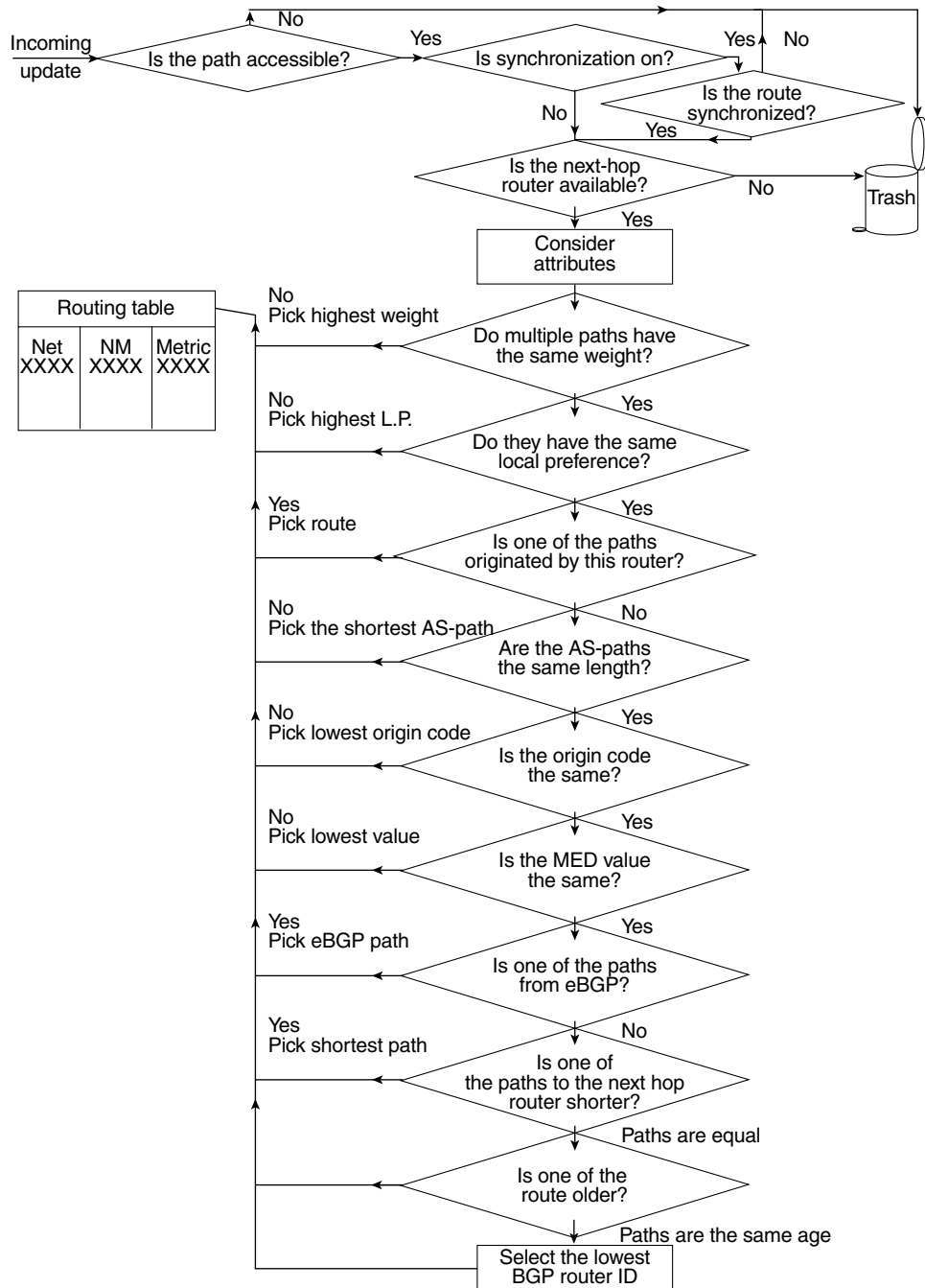
Attribute Name	Category	Code	Preference	Description
Originator ID	Optional, nontransitive	9	Information not used in path selection	The route reflector (described in the following chapter) appends this attribute. It carries the router ID of the originating router in the local autonomous system. It is used to prevent loops.
Cluster ID	Optional, nontransitive	10	Information not used in path selection	The cluster identifies the routers involved in the route reflection. The cluster list shows the reflection path that has been taken. This is used to prevent looping errors.
Weight	Cisco-defined		Highest value	This is proprietary to Cisco and is used in route selection. It is local to the router, and because it is not propagated to other routers, there is no problem with compatibility. When there are multiple paths, it selects a path to a destination with different next hops to the same destination. Note that the weight attribute has no code. Because it is a local attribute and is not propagated to other routers, no code is needed.

Figure 15-10 shows the logic of the path selection used in BGP.

Use Figure 15-10 in association with the following list of the selection process steps:

1. If the router has a valid route to the destination, use that route.
2. If there is more than one valid route to the destination, take the route with the highest weight (Cisco proprietary).
3. If the weights are the same, select the route with the highest local preference.
4. If the routes have the same local preference, prefer the route that originated on the local router.
5. If there are no routes that originated on the router, examine the AS_Path and select the shortest path.
6. If the AS_Path is the same, examine and choose the lowest origin code.
7. If the origin codes are the same, select the path with the lowest MED (the MED values must have been sent from the same neighboring autonomous system).

Figure 15-10 Path Selection in BGP



8. If the MED values are the same, choose an external BGP route over an internal BGP route.
9. If there is no external route, choose the path with the lowest IGP metric or cost to the next-hop router for iBGP.
10. If the paths are equal, select the oldest path.
11. If all else fails, choose the router with the lowest BGP router ID. The router ID is either the highest IP address, with preference given to the loopback interface, or manually configured.

Table 15-7 summarizes the commands covered in this chapter.

Table 15-7 *Summary of BGP Commands*

Command	Function
router bgp <i>autonomous-system-number</i>	Starts the BGP routing process.
network <i>network-number mask network-mask</i>	Identifies the networks to be advertised by the BGP process.
neighbor { <i>ip-address</i> <i>peer-group-name</i> } remote-as <i>autonomous-system-number</i>	Identifies the neighbor with whom the router is synchronizing its routing table and activates a TCP session with the neighbor. It also configures the remote-as option for a peer group.
neighbor { <i>ip-address</i> <i>peer-group-name</i> } next-hop-self	To avoid the problem of selecting the next-hop router on a NBMA network inappropriately, this command is used to force the router to use its own IP address as the next hop when advertising to neighbors.
no synchronization	Turns off synchronization and the need for the IGP to know of a route before BGP can advertise it. This is used when the iBGP network is fully meshed.
aggregate-address <i>ip-address mask</i> [summary-only] [as-set]	Used to create an aggregate address. The summary-only option advertises the summary, and the as-set option lists the autonomous system numbers that the more specific routes have traversed.
debug ip bgp [dampening events keepalives updates]	Enables you to be very specific about the BGP debug parameters.
clear ip bgp { <i>*</i> <i>address</i> } [soft [in out]]	Resets the session between the neighbors and reestablishes it with the new configuration that has been entered. The soft option does not tear down the sessions, but it resends the updates. The in and out options allow the configuration of inbound or outbound soft updates. The default is for both.
show ip bgp [summary neighbors]	Shows the BGP connections. A network can be specified to retrieve information on the lone network. The summary option will give the status of the BGP connections. The neighbors option gives both TCP and BGP connections.

Q&A

As mentioned in the introduction, you have two choices for review questions. The questions that follow next give you a bigger challenge than the exam itself by using an open-ended question format. By reviewing now with this more difficult question format, you can exercise your memory better and prove your conceptual and factual knowledge of this chapter. The answers to these questions are found in Appendix A.

For more practice with examlike question formats, including questions using a router simulator and multichoice questions, use the exam engine on the CD-ROM.

1. If the weight attribute is used, is a higher or lower weight preferred?
2. What is an alternative to using BGP as the method of connection to the ISP?
3. What does the command **clear ip bgp *** achieve, and why should it be used cautiously?
4. Give three situations in which you should not use BGP to connect to the Internet.
5. Explain the use of the command **neighbor 10.10.10.10 remote-as 250**.
6. Explain briefly the purpose of the community attribute.
7. In the route selection process, place the following in order of preference: origin code, weight, local preference, and MED. State the method of selection for the individual attributes themselves.
8. What is a mandatory attribute?
9. What type of routing protocol is BGP classified as, and what does this mean?
10. Explain how static routes can be used as an alternative to configuring BGP to connect to another autonomous system.
11. Explain how a default route can be used as an alternative to configuring BGP to connect to another autonomous system.
12. State two attributes of BGP.
13. State four message types of BGP.
14. What is policy-based routing?
15. What do the letters *MED* represent? Give a brief explanation of what MED does.
16. What is a community in BGP?
17. Give two reasons why BGP peer groups are useful.

18. What is the difference between a peer and a neighbor?
19. In BGP, describe the purpose of the **network** command.
20. Explain the command **neighbor** *{ip-address | peer-group-name}* **next-hop-self** .

Scenarios

The following scenarios and questions are designed to draw together the content of the chapter and to exercise your understanding of the concepts. There is not necessarily a right answer. The thought process and practice in manipulating the concepts are the goals of this section. The answers to the scenario questions are found at the end of this chapter.

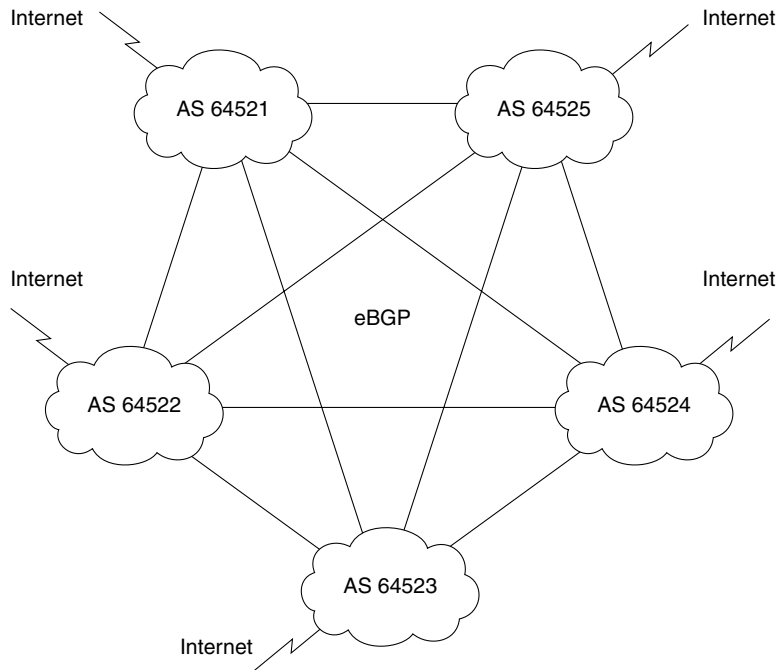
Scenario 15-1

The company Humugos is waiting for the consultant to configure the network. The requirement is to give each country in which Humugos operates its own autonomous system number. The countries will be connected via eBGP and will use leased lines. The autonomous system numbers are private because the connection to the Internet is dealt with by an ISP at each local site. For the first phase of the switchover, EIGRP is removed from the connections between the countries, and the BGP configuration needs to be implemented to ensure a smooth transition. The intention is for each country to have the same configuration to ease management and troubleshooting.

1. Using the diagram in Figure 15-11 as a reference, issue the commands that need to be configured at each country or autonomous system. The private autonomous system numbers range from 64,512 to 65,535.

Figure 15-11 has been simplified and does not contain 250 autonomous systems, as the case study suggests.

2. The BGP network is a full-mesh network. Are there going to be any scaling problems ensuing from this?
3. What commands would indicate that there was a problem of scaling?

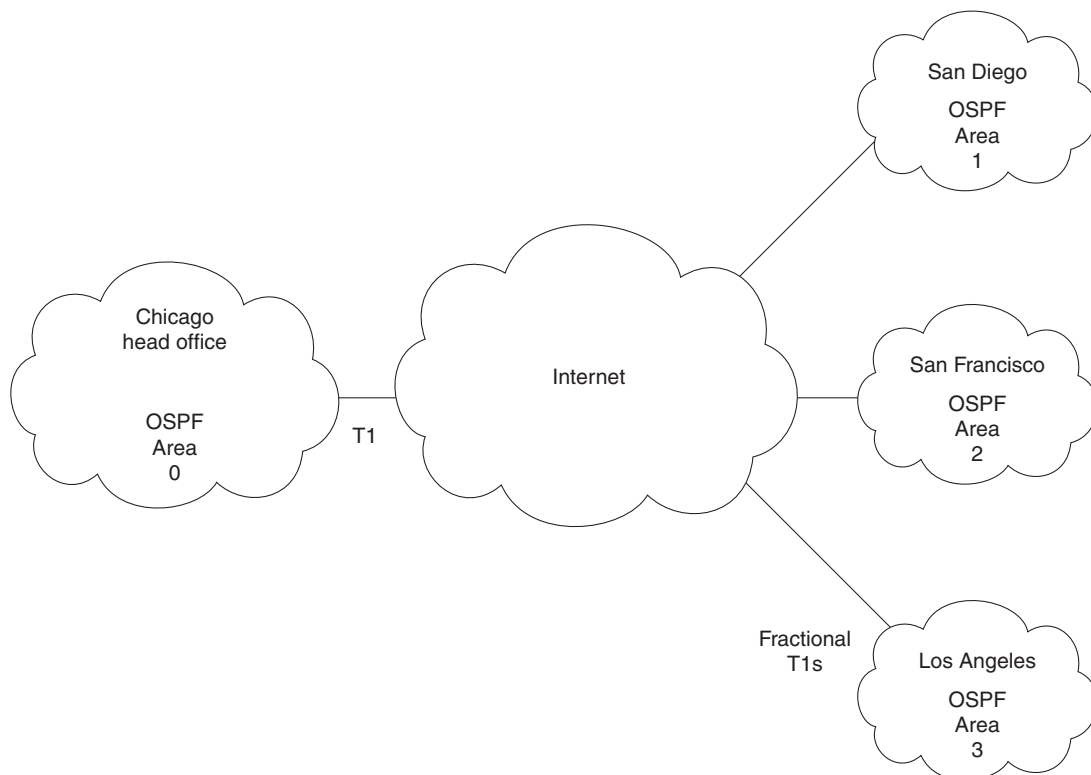
Figure 15-11 *Diagram for Scenario 15-1*

Scenario 15-2

A small company called Insolvent, Inc., has a main office in Chicago and satellite offices on the West Coast of the United States. The company has recently changed its routing protocol to OSPF.

Insolvent has a connection to the Internet from each site, over which it does all its business. The link is a fractional T1 at the satellite offices and a full T1 at the main office. The network administrator at the main office is responsible for the corporate network and is currently trying to recruit staff to manage the local networks. The network administrator was advised at a technical seminar that BGP is what is needed to connect to the Internet. Figure 15-12 shows the network.

1. Given the description of the company and with reference to Figure 15-12, do you agree that BGP is a requirement for this network? Give reasons for your answer.
2. What alternatives are available?
3. Give the alternative configuration commands for the satellite site to connect to Internet.
4. What commands would show that the link is up and operational?

Figure 15-12 *Diagram for Scenario 15-2*

Scenario 15-3

Review the output in Example 15-3, and answer the following questions.

Example 15-3 *Scenario 15-3 Output*

```

FARADAY-gw#show ip bgp neighbor
BGP neighbor is 155.94.83.1, remote AS 2914, external link
Index 1, Offset 0, Mask 0x2
  BGP version 4, remote router ID 129.250.116.16
  BGP state = Established, table version = 457046, up for 1w5d
  Last read 00:00:20, hold time is 180, keepalive interval is 60 seconds
  Minimum time between advertisement runs is 30 seconds
  Received 890723 messages, 0 notifications, 0 in queue
  Sent 36999 messages, 0 notifications, 0 in queue
  Inbound path policy configured
  Outbound path policy configured
  
```

continues

Example 15-3 *Scenario 15-3 Output (Continued)*

```

Outgoing update AS path filter list is 1
Route map for incoming advertisements is limit-verioverdi
Connections established 3; dropped 2
Last reset 1w5d, due to : User reset request
No. of prefix received 11031
No. of prefix received 11031
Connection state is ESTAB, I/O status: 1, unread input bytes: 0
Local host: 155.94.83.2, Local port: 11036
Foreign host: 155.94.83.1, Foreign port: 179
Enqueued packets for retransmit: 0, input: 0 mis-ordered: 0 (0 bytes)
Event Timers (current time is 0x845DFA38):
Timer           Starts      Wakeups      Next
Retrans         18473        11           0x0
TimeWait         0            0           0x0
AckHold         86009        50293        0x0
SendWnd          0            0           0x0
KeepAlive        0            0           0x0
GiveUp           0            0           0x0
PmtuAger         0            0           0x0
DeadWait         0            0           0x0
iss: 829352113 snduna: 829702916 sndnxt: 829702916 sndwnd: 16004
irs: 625978143 rcvnxt: 652708970 rcvwnd: 16342 delrcvwnd: 42
SRTT: 300 ms, RTTO: 607 ms, RTV: 3 ms, KRTT: 0 ms
minRTT: 4 ms, maxRTT: 764 ms, ACK hold: 300 ms
Flags: higher precedence, nagle
Datagrams (max data segment is 536 bytes):
Rcvd: 122915 (out of order: 0), with data: 105023, total data bytes: 26731112
Sent: 109195 (retransmit: 11), with data: 18461, total data bytes: 350802
!
!
BGP neighbor is 144.39.228.49, remote AS 701, external link
Index 2, Offset 0, Mask 0x4
BGP version 4, remote router ID 144.39.3.104
BGP state = Established, table version = 457055, up for 2w0d
Last read 00:00:08, hold time is 180, keepalive interval is 60 seconds
Minimum time between advertisement runs is 30 seconds
Received 50265 messages, 0 notifications, 0 in queue
Sent 37016 messages, 0 notifications, 0 in queue
Inbound path policy configured
Outbound path policy configured
Outgoing update AS path filter list is 1
Route map for incoming advertisements is limit-uunetmemenet
Connections established 2; dropped 1
Last reset 2w0d, due to : Peer closing down the session
No. of prefix received 1635
Connection state is ESTAB, I/O status: 1, unread input bytes: 0
Local host: 144.39.228.50, Local port: 179

```

Example 15-3 *Scenario 15-3 Output (Continued)*

```

Foreign host: 144.39.228.49, Foreign port: 11013
Enqueued packets for retransmit: 0, input: 0 mis-ordered: 0 (0 bytes)
Event Timers (current time is 0x845F16B8):
Timer           Starts      Wakeups          Next
Retrans         20357         4              0x0
TimeWait        0             0              0x0
AckHold         29701         26058          0x0
SendWnd         0             0              0x0
KeepAlive       0             0              0x0
GiveUp          0             0              0x0
PmtuAger        0             0              0x0
DeadWait        0             0              0x0
iss: 3360945234 snduna: 3361331966 sndnxt: 3361331966 sndwnd: 15890
irs: 2976917809 rcvnxt: 2977685910 rcvwnd: 15072 delrcvwnd: 1312
SRTT: 306 ms, RTT0: 642 ms, RTV: 15 ms, KRTT: 0 ms
minRTT: 4 ms, maxRTT: 908 ms, ACK hold: 300 ms
Flags: passive open, nagle, gen tcbs
Datagrams (max data segment is 1460 bytes):
Rcvd: 48675 (out of order: 0), with data: 29705, total data bytes: 768119
Sent: 46955 (retransmit: 4), with data: 20353, total data bytes: 386750

```

1. How many sessions are active?
2. What is the state of the sessions, and what do the states mean?

Scenario Answers

The answers provided in this section are not necessarily the only possible answers to the questions. The questions are designed to test your knowledge and to give practical exercise in certain key areas. This section is intended to test and exercise skills and concepts detailed in the body of this chapter.

If your answer is different, ask yourself whether it follows the tenets explained in the answers provided. Your answer is correct not if it matches the solution provided in the book, but rather if it has included the principles of design laid out in the chapter.

In this way, the testing provided in these scenarios is deeper: It examines not only your knowledge, but also your understanding and ability to apply that knowledge to problems.

If you do not get the correct answer, refer back to the text and review the subject tested. Be certain to also review your notes on the question to ensure that you understand the principles of the subject.

Scenario 15-1 Answers

1. *Using the diagram in Figure 15-11 as a reference, issue the commands that need to be configured at each country or autonomous system.*

The commands configured at each country or autonomous system would be the same structurally, although the details, such as the IP addresses and the autonomous system numbers, would change.

Example 15-4 shows the configuration of the autonomous system 64522, which is the San Francisco router. It has been assigned network 10.2.0.0.

Example 15-4 Scenario 15-1 Configuration

```
Router(config)#
router bgp 64522
no auto-summary
Neighbor 10.1.100.1 remote-as 64521
Neighbor 10.3.100.1 remote-as 64523
Neighbor 10.4.100.1 remote-as 64524
Neighbor 10.5.100.1 remote-as 64525
Neighbor 10.6.100.1 remote-as 64526
Neighbor 10.7.100.1 remote-as 64527
Neighbor 10.8.100.1 remote-as 64528
Neighbor 10.9.100.1 remote-as 64529
Neighbor 10.10.100.1 remote-as 64530
```

Example 15-4 *Scenario 15-1 Configuration (Continued)*

```
!neighbor commands omitted for brevity
Neighbor 10.250.100.1 remote-as 64750
network 10.2.0.0
```

The protocol has had the neighbors in each autonomous system defined with their next-hop IP address and the number of the autonomous system to which they are connecting. The **no auto-summary** command is used to ensure that the subnets of network 10.0.0.0 are advertised; otherwise, each subnet would need to be defined as a **network** command.

2. *The BGP network is a full-mesh network. Are there going to be any scaling problems ensuing from this?*

There should not be a problem with this design. Although there is an enormous number of TCP connections, the traffic is minimal, particularly if aggregation is configured. Also, BGP sends only triggered updates, so if the network is stable and route aggregation is configured, bandwidth should not be a concern. However, considering the propagation delays, the BGP timers might need to be reviewed. Regarding CPU, a high-power router should be used for this purpose. Each time a new BGP neighbor is added, there will be an increasing number of eBGP peer configurations for the network administrator.

3. *What commands would indicate that there was a problem of scaling?*

The commands that should be used to determine whether there is a problem are as follows:

- **show ip bgp neighbors**
- **show ip bgp paths**
- **show ip bgp neighbors**
- **show ip bgp summary**
- **show ip bgp**
- **show processes cpu**

Scenario 15-2 Answers

1. *Given the description of the company and with reference to Figure 15-12, do you agree that BGP is a requirement for this network? Give reasons for your answer.*

Because the company is small and has only a single connection per site into the Internet, it would be too complex to configure and maintain BGP when there simply are not enough resources. The bandwidth is inadequate for the task, and the administrative expertise is already overstretched. It would be far better to configure one or two static/default routes to the Internet and to redistribute these routes into the IGP running within the autonomous system.

2. *What alternatives are available?*

The only real alternative is the one already mentioned: to configure a default route into the Internet from every location and to redistribute this default route into the IGP for the autonomous system.

3. *Give the alternative configuration commands for the satellite site to connect to Internet.*

Each site would have the same configuration structure, although the details might differ:

```
ip route 0.0.0.0 0.0.0.0 155.94.83.1
router ospf 100
network 207.111.9.0 0.0.0.255 area 0
default-information originate always
```

The first line configures the default route. The second line turns on the OSPF process 100. The third line identifies which interfaces are participating in OSPF and what area they are in. The fourth line propagates the default route into the network, whether or not the advertising router has a path to the network.

4. *What commands would show that the link is up and operational?*

The commands to prove that the link is up and operational would be the **show ip route** command and the **ping** and **traceroute** commands. Refer to the Cisco web site or the ICND course for more details on these commands.

Scenario 15-3 Answers

1. *How many sessions are active?*

There are two active sessions. In reading the large amount of information on the **show ip bgp neighbor** command, there is a line at the beginning of each session identifying the neighboring peer. The lines in this output screen are as follows:

```
BGP neighbor is 155.94.83.1, remote AS 2914, external link
BGP neighbor is 144.39.228.49, remote AS 701, external link
```

2. *What is the state of the sessions, and what do the states indicate?*

```
BGP state = Established, table version = 457046, up for 1w5d
BGP state = Established, table version = 457055, up for 2w0d
```

Both the peers have established sessions. This means that they have a TCP session between them. They are now in a position to exchange routing tables and to synchronize their databases. The rest of the line indicates how many times the table has been updated and how long the session has been maintained. In this example, the first peer has had a session with the local router for one week and five days, while the second peer has been up for exactly two weeks.

There are three autonomous systems in this configuration. The first peer belongs to autonomous system 2914, and the second belongs to 701. Because both of these neighbors belonging to their autonomous systems have an external BGP session, there must be a third autonomous system, within which the local router resides.

```
BGP neighbor is 155.94.83.1, remote AS 2914, external link
BGP neighbor is 144.39.228.49, remote AS 701, external link
```



This chapter covers the following topics, which you need to understand to pass the CCNP/CCDP/CCIP BSCI exam:

- Building a network using internal BGP
- Understanding iBGP requirements
- Designing and configuring a scalable iBGP network
- Verifying the iBGP configuration
- Controlling BGP traffic
- Connecting to the Internet with BGP
- Determining the BGP path by tuning the attributes
- Redistribution between IGP and BGP

Implementing and Tuning BGP for Use in Large Networks

The topics in this chapter concern the advanced configuration of BGP. Whereas Chapter 15, “Connecting to Other Autonomous Systems—The Basics of BGP,” discussed basic concepts and configuration of BGP, this chapter delves into some of the complexities of BGP. In this chapter, you explore the uses of BGP—whether connecting to an ISP or even acting as an ISP with several connected organizations. The chapter also deals with the use of internal BGP (iBGP), which is BGP configured for systems within an autonomous system. The chapter covers how BGP can be configured to select a particular path and the design features and pitfalls of BGP. In this discussion of the advanced configuration of BGP, the explanation of the technology is coupled with configuration examples so that your conceptual understanding of BGP is reinforced by concrete implementation examples.

“Do I Know This Already?” Quiz

The purpose of the “Do I Know This Already?” quiz is to help you to decide what parts of this chapter to use. If you already intend to read the entire chapter, you do not necessarily need to answer these questions now.

The 18-question quiz, derived from the major sections in the “Foundation Topics” portion of the chapter, helps you to determine how to spend your limited study time.

Table 16-1 outlines the major topics discussed in this chapter and the “Do I Know This Already?” quiz questions that correspond to those topics.

Table 16-1 “Do I Know This Already?” Foundation Topics Section-to-Question Mapping

Foundation Topics Section	Questions Covered in This Section
Building a Network Using Internal BGP	1
Understanding iBGP Network Requirements	2–4
Designing and Configuring a Scalable iBGP Network	5–7
Verifying the iBGP Configuration	8–9

(continues)

Table 16-1 “Do I Know This Already?” Foundation Topics Section-to-Question Mapping (Continued)

Foundation Topics Section	Questions Covered in This Section
Controlling BGP Traffic	10–12
Connecting to the Internet with BGP	13
Determining the BGP Path by Tuning the Attributes	14–16
Redistribution Between IGP and BGP	17–18

CAUTION The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark this question wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

1. How does the router determine whether it has received an iBGP or an eBGP update?
 - a. The router is configured to iBGP or eBGP; it cannot determine between the two.
 - b. The iBGP updates are not propagated; a received update must be from eBGP.
 - c. The router sees whether the received update has routes in the IGP routing table.
 - d. The router checks the autonomous system number in the open message that was sent.
2. Which of the following is true of iBGP neighbors?
 - a. iBGP does not require neighbors to be physically adjacent.
 - b. iBGP requires the neighbors to be physically adjacent.
 - c. iBGP neighbors must share the same IP subnet address.
 - d. iBGP neighbors must be used on a broadcast medium.
3. When is it possible to turn off the default synchronization?
 - a. When confederations are used
 - b. When route reflectors are configured
 - c. When the iBGP neighbors are fully meshed
 - d. When all the routers in the autonomous system are running iBGP
4. Why is it necessary to have a fully meshed iBGP network?
 - a. No updates are propagated by other clients.

- b. It is not possible to redistribute between the IGP and BGP.
 - c. To prevent routing loops.
 - d. BGP treats routing within the autonomous system as if it were an NBMA network.
- 5. Which of the following is true of route reflectors?
 - a. Route reflectors disallow the forwarding of attributes.
 - b. Route reflectors prevent the need for a fully meshed eBGP network.
 - c. Route reflectors are used to reflect routes to all neighbors.
 - d. Route reflectors forward routing updates to neighbors or within the autonomous system.
- 6. Which of the following is true of a route reflector client?
 - a. A client is a router that sends updates to the route reflector to be forwarded to other clients.
 - b. A client is a router that receives updates from a route reflector.
 - c. A client is a router that is configured and acts as a stub router.
 - d. A client is a router that receives routes from an IGP.
- 7. Select the statement that is true.
 - a. The routing table for IP contains the selected BGP path and the attributes.
 - b. BGP uses the command **clear ip bgp *** to clear the IP routing table.
 - c. By default, BGP does not store prefixes before policy application.
 - d. BGP sends incremental updates every 15 minutes.
- 8. Which of the following fields are shown in the **show ip bgp neighbors** command?
 - a. Neighbors
 - b. The prefixes associated with each neighbor
 - c. State of the BGP connections
 - d. The last time a message was read from a neighbor
- 9. What command is used to display the route reflector clients?
 - a. **show ip ibgp neighbors**
 - b. **show ip bgp neighbor**
 - c. **show ip bgp route**
 - d. **show ip route reflector**

10. Which of the following filtering techniques can be used to filter on prefixes?
 - a. AS_Path lists
 - b. Distribute lists
 - c. Access lists
 - d. Prefix lists

11. When using a prefix list, what is the match criterion used?
 - a. The first match
 - b. The attribute name
 - c. The match between **ge** and **le**
 - d. All matches

12. Which command is used to configure a prefix list?
 - a. **ip bgp prefix list**
 - b. **ip prefix-list**
 - c. **access-list (200-299)**
 - d. **bgp prefix list (200-299)**

13. What are the most common forms of multihoming?
 - a. ISPs send default routes into the autonomous system.
 - b. Complete routing tables are sent into the IGP.
 - c. Default and some selected routes are sent by the ISP into the autonomous system.
 - d. All routes known by ISP are sent into the autonomous system.

14. In the command **show ip bgp** , the origin field is placed at the end of each line in the table. It can be which of the following values?
 - a. Entry originated from IGP and was advertised with a network router configuration command.
 - b. Entry originated from an EGP.
 - c. Origin of the path is not clear. Usually, this is a router that is redistributed into BGP from an IGP.
 - d. The IP address of the originating router.

15. Which command shows the local preference and weight attribute values?
 - a. **show bgp attributes**
 - b. **show ip bgp**
 - c. **show ip bgp path**
 - d. **show ip bgp attributes**

16. Give the command that would change the weight attribute for the path to the next hop 192.10.5.3.
 - a. **set weight 192.10.5.3 48**
 - b. **neighbor 192.10.5.3 weight 48**
 - c. **set attribute weight 48**
 - d. **ip bgp neighbor 192.10.0.0 remote-as 300 weight 48**

17. Which are the ways to advertise routes into BGP?
 - a. They are automatically redistributed
 - b. The **network** command
 - c. Redistribute static routes
 - d. Redistribute dynamically learned routes from the IGP

18. What is the concern when advertising routes from BGP into an IGP?
 - a. BGP will receive too many routes and increase the size of the Internet routing tables.
 - b. Private addresses will be propagated into the Internet.
 - c. Routing loops will occur.
 - d. The IGP will be overwhelmed.

The answers to the "Do I Know This Already?" quiz are found in Appendix A, "Answers to Chapter 'Do I Know This Already?' Quizzes and Q&A Sections." The suggested choices for your next step are as follows:

- **9 or less overall score**—Read the entire chapter. This includes the "Foundation Topics" and "Foundation Summary" sections, the "Q&A" section, and the "Scenarios" at the end of the chapter.
- **10–14 overall score** —Begin with the "Foundation Summary" section, and then go to the "Q&A" section and the "Scenarios" at the end of the chapter. If you have trouble with these exercises, read the appropriate sections in "Foundation Topics."
- **15 or more overall score** —If you want more review on these topics, skip to the "Foundation Summary" section, and then go to the "Q&A" section and the "Scenarios" at the end of the chapter. Otherwise, move to the next chapter.

Foundation Topics

Building a Network Using Internal BGP

Though BGP is an exterior routing protocol, it comes in two flavors: internal BGP (iBGP) and external BGP (eBGP). The difference depends on the function of the routing protocol. The router will determine whether the peer BGP router is going to be an eBGP peer or an iBGP peer by checking the autonomous system number in the open message that was sent.

eBGP complies with the common perception of an external routing protocol; it sends routing information between different autonomous systems. Therefore, the border router between different autonomous systems is the eBGP router.

iBGP is used within an autonomous system. It conveys information to all BGP routers within the domain and ensures that they have a consistent understanding of the networks available.

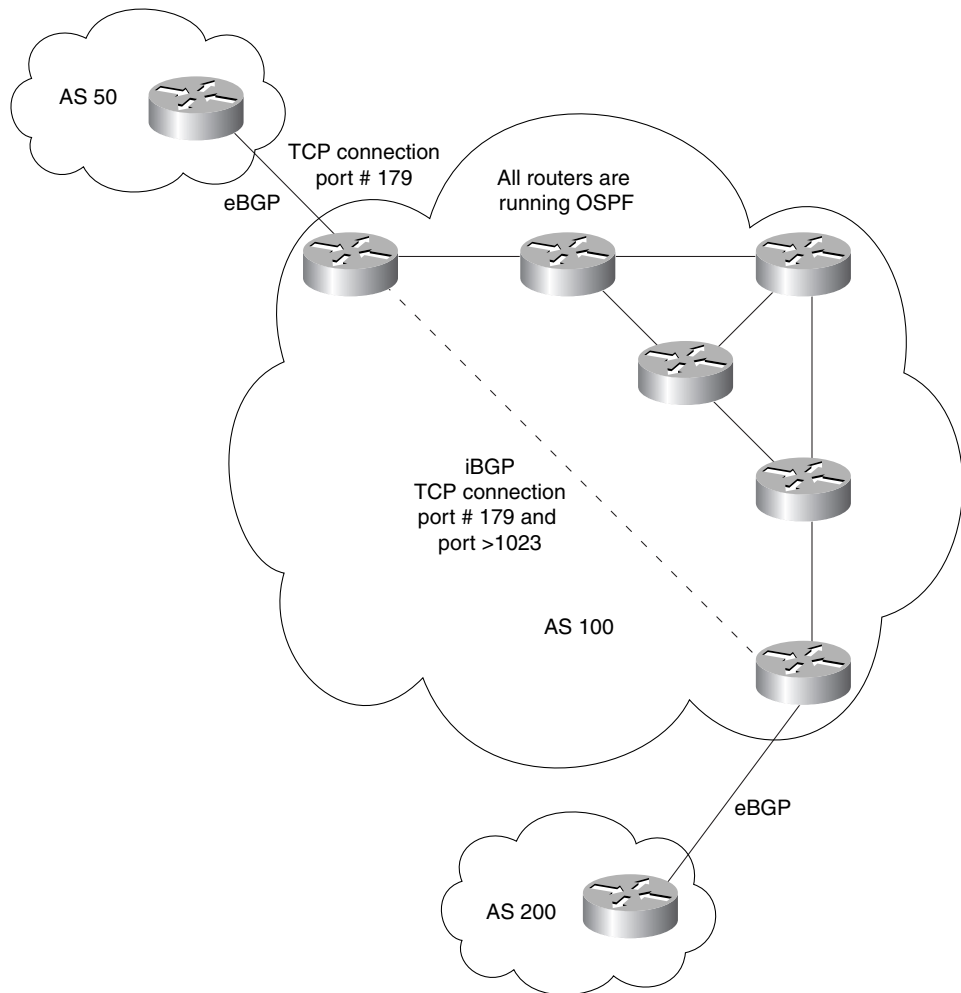
iBGP is used by an ISP or a large organization when it is necessary to convey external BGP information about other autonomous systems across a *transit autonomous system*; that is, iBGP is used between eBGP speakers in the same autonomous system. In Figure 16-1, eBGP is used between AS 50 and AS 100, but in order to connect to AS 200, the BGP routing information must traverse AS 100. Once the routing protocol has traveled through AS 100, it returns to its more natural form of eBGP to connect to AS 200.

Understanding iBGP Network Requirements

To design and implement BGP correctly, there are a few characteristics about iBGP that are important to understand. The first of these rules, discussed in the next section, states that iBGP routers do not need to be directly connected, whereas eBGP routers must be physically connected (unless they are running multihop eBGP). You then learn about how synchronization and fully meshed networks factor into BGP network design.

Physical versus Logical Connections

Unlike Internal Gateway Protocols (IGPs), such as RIP, OSPF, and IPv6, that carry information about the autonomous system, the iBGP routers are not required to be physical neighbors on the same medium. In fact, they often sit at the edges of the autonomous system. Another routing protocol, an IGP such as OSPF, routes the BGP packets between the iBGP routers. In Figure 16-1, you see the iBGP routers are not directly connected, but that they have a TCP connection using TCP port number 179. This means that the internal protocol has topology independence.

Figure 16-1 *iBGP and eBGP*

Through a logical, not a physical, connection, TCP routes the BGP packets across the autonomous system by routers with routing tables maintained by OSPF. For BGP to communicate the routing information, it redistributes its routing information into the IGP. The integration of these different routing protocols can be challenging.

Now that you understand the topological requirements of iBGP, you can follow the criteria by which iBGP sends updates over this topology. An iBGP router propagates a route to a BGP neighbor under the following conditions:

- If the advertised route was generated by the transmitting router by one of the following methods:
 - Via the **network** command
 - Redistributed from an IGP
 - Redistributed static routes
- If the advertised route is a connected route

These criteria are important to understand as they necessitate some design restrictions. Essentially, if the route was learned via an update from a BGP peer within the same autonomous system, a BGP router can propagate this route only to an eBGP peer.

Because iBGP does not forward updates it learned from an iBGP peer, iBGP peers need to be connected to one another (fully meshed) to have a complete knowledge of the network. However, a fully meshed network makes it difficult to scale iBGP because every router in the autonomous system has to build a BGP session with every other router in the autonomous system.

Synchronization

A simple rule states that before iBGP can propagate a route into another autonomous system by handing it over to eBGP, the route must be known throughout the autonomous system. That is to say, the IGP or internal routing protocol must be synchronized with BGP.

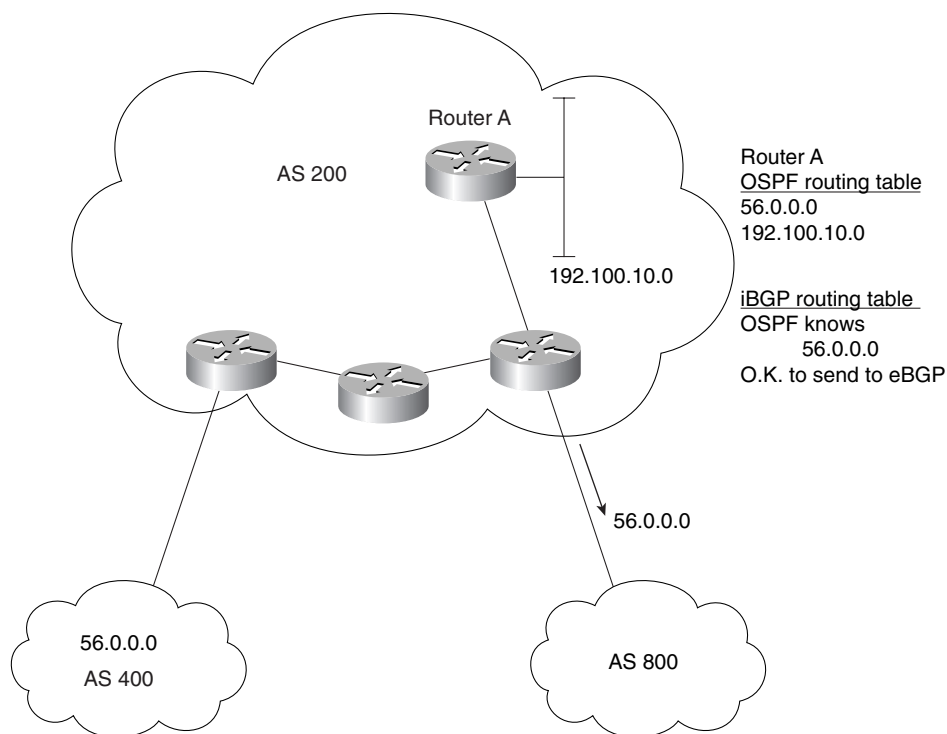
This synchronization is to ensure that if traffic is sent into the autonomous system, the interior routing protocol can direct it to its destination. The synchronization rule is on by default and should be turned off only if all routers in the autonomous system are running BGP.

For example, if you have a transit autonomous system with only the edge routers running iBGP, you are relying on the IGP to carry the traffic between the iBGP routers. It is therefore important for the IGP to have the information in its routing table to fulfill this task. This example is illustrated in Figure 16-2.

As you can see, AS 400 and AS 800 are using AS 200 as a transit autonomous system. In accordance with the synchronization rule, the router sending updates into AS 800 will not propagate 56.0.0.0 advertised by AS 400 unless the IGP, in this case OSPF, has full knowledge of how to get to 56.0.0.0. Otherwise, AS 800 might send traffic to 56.0.0.0, which would get lost in AS 200, because with an incomplete IP routing table, it would not be able to direct the traffic to the appropriate destination.

The synchronization rule is beneficial for the following reasons:

- It prevents traffic from being forwarded to unreachable destinations.
- It reduces unnecessary traffic.
- It ensures consistency within the autonomous system.

Figure 16-2 *Synchronization Rule and BGP*

On some occasions, it is useful to turn off synchronization. As with any default, it is unwise to turn off this option without a detailed understanding of the network. The occasions when it might be useful to turn off synchronization are as follows:

- If all the routers in the autonomous system are running BGP.
- If all the routers inside the autonomous system are meshed.
- When the autonomous system is not a transit autonomous system.

To turn off synchronization, you can use the following command in config-router mode:

```
Router(config-router)#no synchronization
```

This allows routers to advertise routes into BGP before the IGP has a copy of the route in its routing table.

A Fully Meshed Network

The BGP split horizon rule states that though the routers do not need to be directly connected, they do need to be fully meshed. This configuration is required to ensure full connectivity. To avoid routing loops, the protocol must follow the split horizon rule that no updates learned from internal peers can be sent to other internal peers. This means that the prefixes it receives are not redistributed to other iBGP systems. The prefixes are only propagated to a BGP system in another autonomous system, otherwise known as a *peer eBGP system*.

Network Resources Required in Fully Meshed Networks

BGP maintains up-to-date and accurate routing information by sending incremental updates across a TCP connection. The TCP connection is an excellent way of ensuring the accuracy of the information, but it is costly in network resources. The greater the number of connections, the greater the number of required resources. A simple equation demonstrates the problem, one that is also seen in the consideration of designing fully meshed Frame Relay networks.

The equation for determining the number of sessions required is as follows:

$$n(n - 1) / 2$$

In plain English, it is the number of routers minus 1, multiplied by the number of routers, and then divided by 2. Thus, 10 routers would mean $10(10 - 1) / 2 = 10 * 9 / 2 = 45$ sessions.

This equation works well in environments that require a few connections, such as a company with multiple connections into the Internet. However, if the network is an ISP that is using BGP throughout its network, some careful design should be put in place.

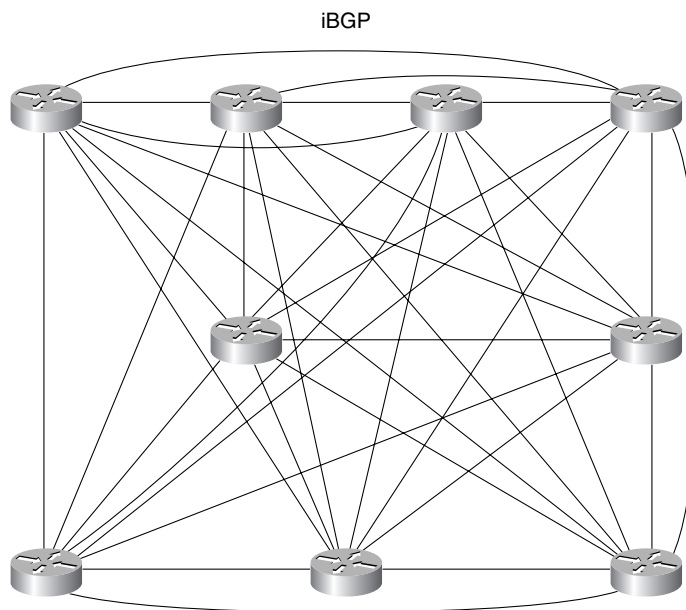
Administrative Overhead in Fully Meshed Networks

There is also administrative overhead in maintaining a fully meshed network of peers. For example, each time a peer is added, the number of iBGP peering configuration statements grows as well.

To be fair to TCP, it is not simply the maintenance of the connection that eats up resources, but the amount of updates that traverse those links. If every router is connected to every other router, a great deal of the information that is being sent is duplicated. Figure 16-3 illustrates the redundancy.

Designing and Configuring a Scalable iBGP Network

The problem of scalability presented by a fully meshed iBGP network can be solved by design. Configuration solutions allow you to overcome the rule that all iBGP peers must be fully meshed. These new commands allow you to develop a hub-and-spoke network to streamline the TCP connections. This is a good thing, but it does require some additional design and configuration. The solution is the implementation of route reflectors and the network design that they support.

Figure 16-3 *iBGP and a Fully Meshed Network*

The following section covers the use of route reflectors that reflect the updates from one connected client to other clients connected to the route reflector. The other topics discussed the next sections include route refresh, which resets the TCP connections and thus affects the change in configuration, and the use of peer groups. To streamline not only the configuration, but also the BGP traffic necessary to maintain full and accurate tables, peer groups are formed to allow one update to be sent to an identified group.

Route Reflectors

A *route reflector* is a router configured to forward routing updates to neighbors or peers within the same autonomous system. These iBGP peers need to be identified as clients in the configuration. When a client sends an update to the route reflector, it is forwarded or reflected to the other clients. Essentially, the route reflector defies the split horizon rule that states that the iBGP router will not propagate a route that was learned from a peer within the same autonomous system (an iBGP peer).

A *client* is a router that receives updates from a route reflector that a route reflector has forwarded from another client. It requires no configuration and has no idea that it is anything other than a peer.

The route reflector and the client require a full peer relationship because the route reflector forwards updates from other clients, but peering between the clients is not needed.

In all probability, a route reflector is connected to peers for which it is not forwarding routes. From the route reflector's view, these neighbors or peers are *nonclients*. Nonclients must be fully meshed with the route reflector and each other.

When a router has been configured as a route reflector, it forwards paths learned from iBGP peers only to route reflector clients and to iBGP/eBGP neighbors. This means that a logical hub-and-spoke design can be implemented within an autonomous system between iBGP peers, thus reducing the number of peering sessions required.

Both a route reflector and its clients form a unit that shares information. This unit is called a *cluster*. The autonomous system can be divided into clusters, and the router reflectors identified and configured. There must be at least one route reflector per cluster. If a route reflector connects to other route reflectors, the route reflectors should be fully meshed. This is to ensure that the iBGP routing tables are complete.

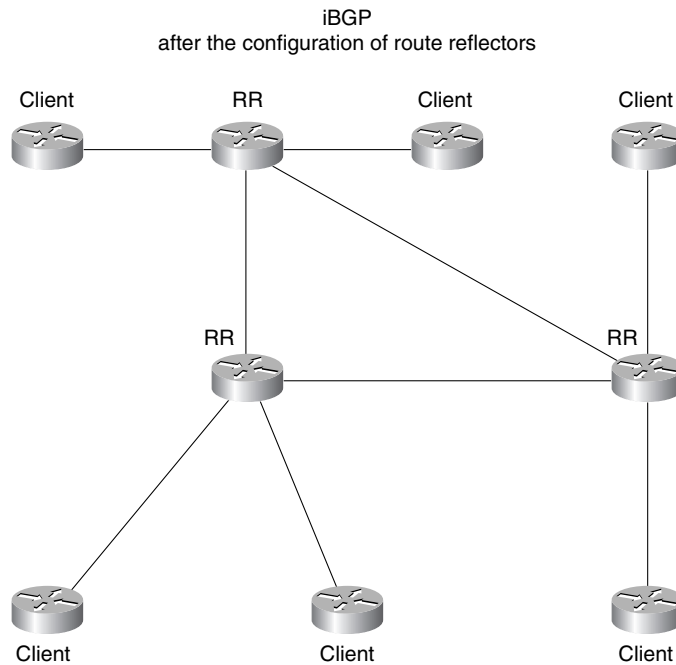
When the route reflector forwards an update, the *Originator-ID* attribute is set. This is the BGP router ID of the router that originated the path. If this router receives back the update, it will see its own ID and will ignore the packet. This prevents the possibility of routing loops.

If there are multiple route reflectors in the cluster to provide redundancy, the originating router is identified by the *Cluster-ID* attribute. This serves the same purpose as the Originator-ID in preventing routing loops.

The route reflector concept means that there is more overhead on the route reflector, and if it is configured incorrectly, it can cause serious routing loops. The design to avoid a fully meshed iBGP network can become quite complicated, but multiple route reflectors afford redundancy, which is always reassuring. Multiple levels of route reflectors can even be configured, creating a hierarchical design.

Nonroute reflector routers are not affected by the change in design and routing update propagation. Indeed, they are blissfully unaware of any changes because they still receive the updates that they need. The updates are also unchanged because no changes are made to the attribute values. This makes migration to a network design incorporating route reflectors very straightforward.

An important design suggestion is for the iBGP route reflectors to be fully meshed to ensure the correct propagation of updates. As mentioned earlier, it is possible to create a hierarchical design where route reflectors are clients of other route reflectors. This is a complex design and requires great care, because as soon as the route reflector is configured and split horizon disabled, there is no protection against a routing loop. A fully meshed route reflector design is therefore advised, as illustrated in Figure 16-4.

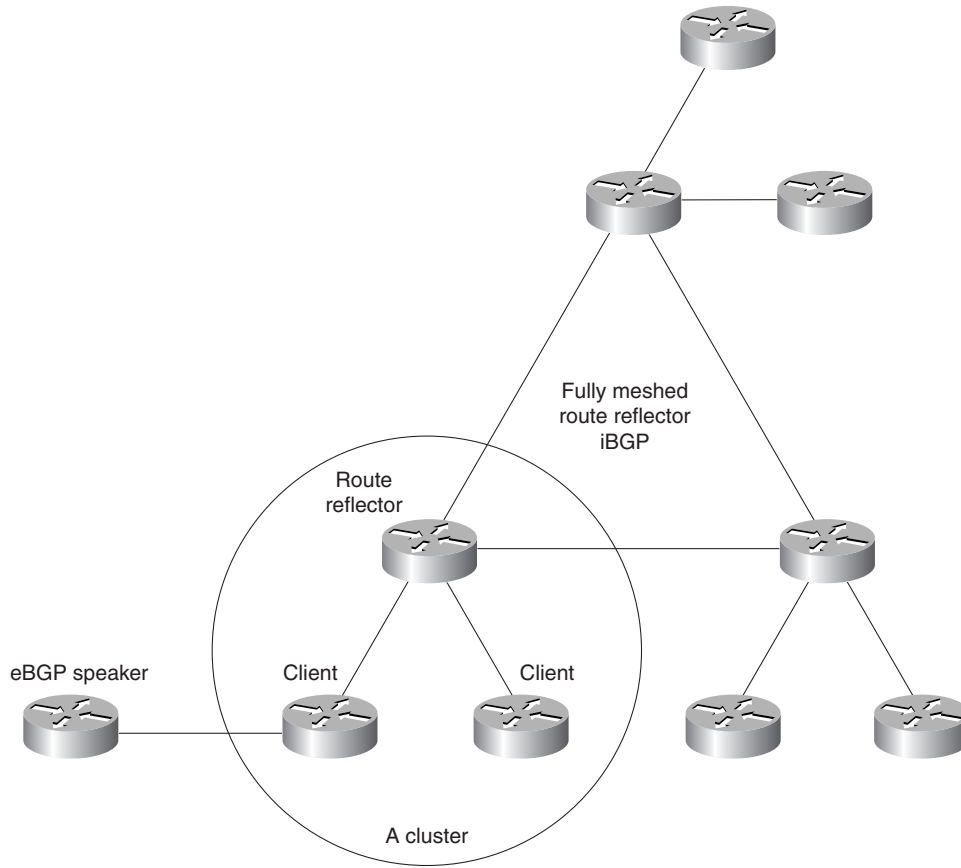
Figure 16-4 *Design of iBGP Network Using Route Reflectors*

The benefits of route reflectors include the following:

- The capability to scale the network
- A strong hierarchical design
- A reduction of traffic on the network
- A reduction in the memory and CPU needed to maintain TCP sessions on the client iBGP peers
- Faster convergence and a simpler network because two routing protocols are implemented:
 - iBGP for external routing information traversing the autonomous system
 - IGP for routes internal to the autonomous system

The solution provided by route reflectors is used in large iBGP environments such as ISP networks, where a fully meshed iBGP network could result in a large number of TCP sessions. Figure 16-5 illustrates the relationship between route reflectors, clients, and other clusters.

The next sections examine how route reflectors operate, how the clients of route reflectors operate, and introduces the concept of a cluster in BGP.

Figure 16-5 *Clusters and Route Reflector Meshing*

How Route Reflectors Operate

The previous section summarized of some of the characteristics of route reflectors. These characteristics determine how route reflectors operate. A route reflector is simply a mirror that reflects updates from clients to other clients without requiring a fully meshed network.

The following list shows what happens when a route reflector receives an update:

- The client forwards an update to its peer, in this case, the route reflector.
- An update from a client is received by the route reflector, and the update is forwarded to other clients as well as nonclients (both iBGP and eBGP peers). The originator ID is excluded from the update.
- If multiple paths are received by the route reflector, the best path is chosen by the route reflector.

- The only router that does have the update forwarded to it is the originator of the route.
- A nonclient forwards an update to its peer, which happens to be a route reflector.
- The route reflector receives an update from a nonclient, and the update is forwarded only to clients.
- An eBGP peer sends an update to the route reflector.
- The route reflector reflects the update to both clients and nonclients.

Configuring Route Reflectors

The command for configuring a route reflector is very straightforward. It is explained in the following syntax:

```
Router(config-router)#neighbor ip-address route-reflector-client
```

To remove a router as a client, issue the following command:

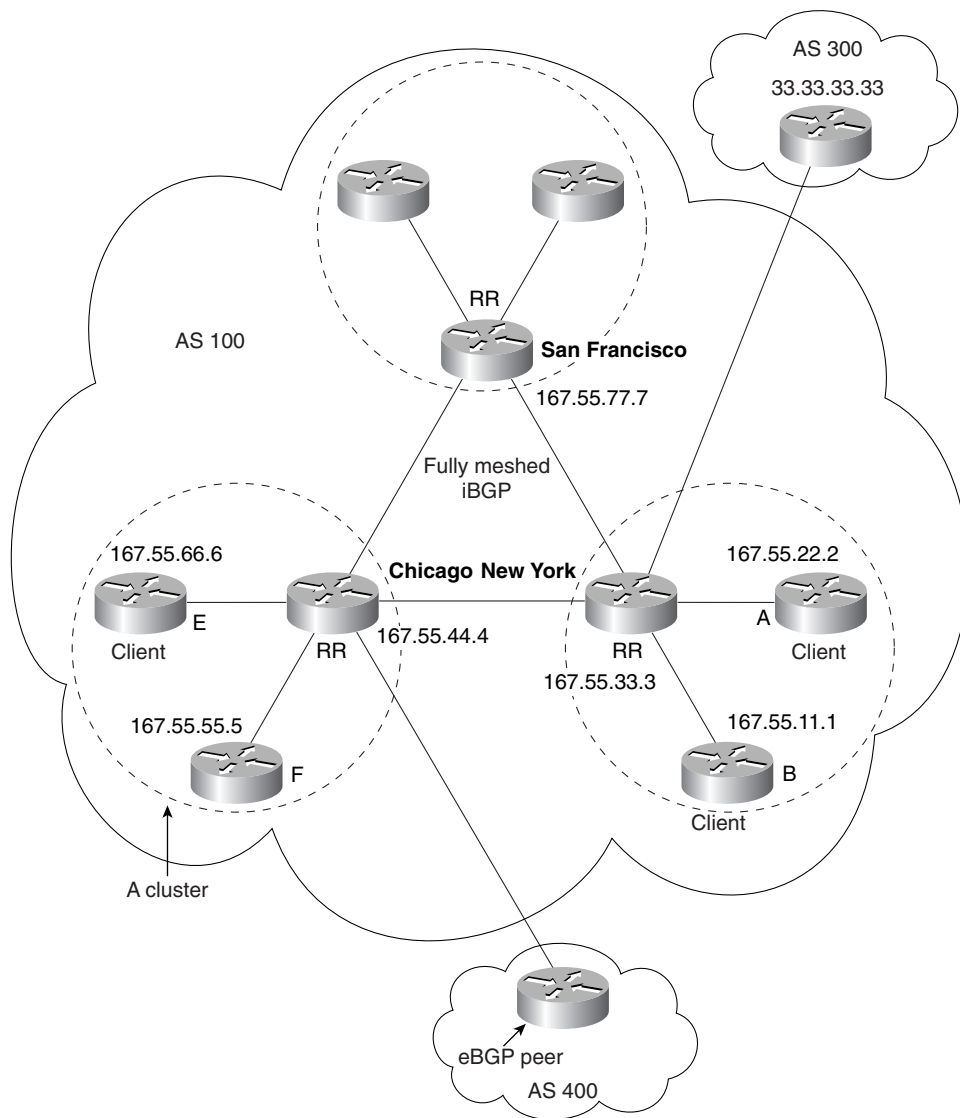
```
Router(config-router)#no neighbor ip-address route-reflector-client
```

Remember that if all clients are removed, the route reflector loses its status and becomes a standard iBGP router. If this happens, the iBGP routers need to be fully meshed. Table 16-2 breaks down the syntax of the command to configure a route reflector and identify the clients.

Table 16-2 *Explanation of the Route Reflector Configuration Command*

Syntax	Description
neighbor	Identifies that the rest of the command is directed at a BGP peer.
<i>ip-address</i>	The IP address of the neighboring router being identified as a client.
route-reflector-client	Points to the client of the route reflector. Note that the client is not configured and is unaware of its change of status. It does nothing but continues to send updates to the route reflector, which forwards them unchanged to other clients.

Example 16-1 illustrates the concepts explained in this section. For simplicity, the connection to the eBGP router in AS 400 has not been included in the example. Use this example in conjunction with the network displayed in Figure 16-6.

Figure 16-6 *Network Topology Configured in Example 16-1***Example 16-1** *Configuration of a Route Reflector*

```

NewYork#
NewYork(config)#router bgp 100
! Create a connection to Router A, a client
NewYork(config-router)#neighbor 167.55.22.2 remote-as 100
! make Router A as seen in Figure 16-6 a route reflector client

```


Example 16-1 *Configuration of a Route Reflector (Continued)*

```

New York(config-router)#neighbor 167.55.22.2 route-reflector-client
! connection to Router B
New York(config-router)#neighbor 167.55.11.1 remote-as 100
New York(config-router)#neighbor 167.55.11.1 route-reflector-client
! connection to San Francisco
New York(config-router)#neighbor 167.55.77.7 remote-as 100
! connection to Chicago
New York(config-router)#neighbor 167.55.44.4 remote-as 100
New York(config-router)#neighbor 33.33.33.33 remote-as 300
RouterB#
RouterB(config)#router bgp 100
! connection to New York
RouterB(config-router)#neighbor 167.55.33.3 remote-as 100
Chicago#
Chicago(config)#router bgp 100
! connection to Router E
Chicago(config-router)#neighbor 167.55.66.6 remote-as 100
Chicago(config-router)#neighbor 167.55.66.6 route-reflector-client
! connection to Router F
Chicago(config-router)#neighbor 167.55.55.5 remote-as 100
Chicago(config-router)#neighbor 167.55.55.5 route-reflector-client
! connection to San Francisco
Chicago(config-router)#neighbor 167.55.77.7 remote-as 100
! connection to New York
Chicago(config-router)#neighbor 167.55.33.3 remote-as 100

```

Route Refresh

After any BGP configuration, it is necessary to reset the TCP session so that the changes can take effect. This is because the BGP process stores only prefixes that apply to the stated policy. If the policy changes, which means after any configuration, the peer session is torn down and rebuilt with the new characteristics.

It is now possible to issue a soft process reboot, which still destroys and rebuilds the peering sessions, but without a hard reboot of the BGP process.

The command to reboot all sessions is as follows:

```
Router#clear ip bgp *
```

The command to tell the peer to resend a full BGP update to a particular neighbor follows:

```
Router#clear ip bgp neighbor-address in
```

The command to tell the process to send a full BGP update to the peer follows:

```
Router# clear ip bgp neighbor-address out
```

The **clear ip bgp** command is described in detail in Chapter 15.

It is also possible to configure the BGP process to store the prefixes before the policy application. This obviously requires greater memory, but it allows new configurations to be implemented without tearing down peering sessions.

The configuration is applied on a per-neighbor basis and only needs to be applied to the inbound updates. The syntax is as follows:

```
Router(config-router)#neighbor neighbor-address soft-configuration inbound
```

After a configuration change has been made, issue the following command from the executive level:

```
Router# clear ip bgp neighbor-address soft [in|out]
```

Peer Groups

Without peer groups, every iBGP peer—being fully meshed—receives the same update. This means every iBGP router performs the same calculations, wasting CPU and restricting the ability of iBGP to scale.

Once peer groups are configured, every router within the peer group has the same outbound policy, while allowing different inbound policies to be configured on each system. This means that one update can be generated for the group, resulting in the following benefits:

- The administrative overhead is reduced, because the configuration is simpler, reducing the possibility of errors.
- Less CPU is required, speeding up the network responsiveness. When a network converges quickly, it becomes more stable and reliable.

To define a peer group, use the following:

```
Router(config-router)#neighbor peer-group-name peer-group
```

Example 16-2 shows how the peer group IBGP-peergp is created and applied to iBGP neighbors.

Example 16-2 Configuration of a Peer Group to iBGP Neighbors

```
Router(config)#router bgp 100
Router(config-router)#neighbor IBGP-peergp peer-group
Router(config-router)#neighbor IBGP-peergp remote-as 100
Router(config-router)#neighbor IBGP-peergp route-map setpolicy out
Router(config-router)#neighbor 6.6.6.6 peer-group IBGP-peergp
Router(config-router)#neighbor 6.7.7.7 peer-group IBGP-peergp
Router(config-router)#neighbor 6.8.8.8 peer-group IBGP-peergp
```

Verifying the iBGP Configuration

It is also important to verify that a configuration is working. Example 16-3 demonstrates the command that verifies the configuration of router 167.55.44.3 and its neighbors and that inbound soft reconfiguration was configured and works.

Example 16-3 *Example of the show ip bgp neighbors Command*

```
New York# show ip bgp neighbors 167.55.44.4
BGP neighbor is 167.55.44.4, remote AS 100, Internal link
Index 1, Offset 0, Mask 0x2
Inbound soft reconfiguration allowed
BGP version 4, remote router ID 167.55.44.4
BGP state = Established, table version = 27, up for 00:06:12
Last read 00:00:12, hold time is 180, keepalive interval is 60 seconds
Minimum time between advertisement runs is 30 seconds
Received 19 messages, 0 notifications, 0 in queue
Sent 17 messages, 0 notifications, 0 in queue
Inbound path policy configured
Route map for incoming advertisements is testing
Connections established 2; dropped 1
Connection state is ESTAB, I/O status: 1, unread input bytes: 0
Local host: 167.55.44.3, Local port: 11002
Foreign host: 167.55.44.4, Foreign port: 179

Enqueued packets for retransmit: 0, input: 0, saved: 0
Event Timers (current time is 0x530C294):
Timer           Starts    Wakeups      Next
Retrans           12         0           0x0
TimeWait          0          0           0x0
AckHold          12         10           0x0
SendWnd           0          0           0x0
KeepAlive         0          0           0x0
GiveUp            0          0           0x0
PmtuAger          0          0           0x0
iss: 133981889 snduna: 133982166 sndnxt: 133982166 sndwnd: 16108
irs: 3317025518 rcvnxt: 3317025810 rcvwnd: 16093 delrcvwnd: 291
SRTT: 441 ms, RTT0: 2784 ms, RTV: 951 ms, KRTT: 0 ms
minRTT: 0 ms, maxRTT: 300 ms, ACK hold: 300 ms
Flags: higher precedence, nagle
Datagrams (max data segment is 1460 bytes):
Rcvd: 15 (out of order: 0), with data: 12, total data bytes: 291
Sent: 23 (retransmit: 0), with data: 11, total data bytes: 276
```

Table 16-3 describes the fields shown in Example 16-3.

Table 16-3 *Explanation of the show ip bgp neighbors Command*

Field	Descriptions
BGP neighbor	IP address and autonomous system of the BGP neighbor. If the autonomous system numbers are the same, then iBGP is running between the neighbors; otherwise, eBGP is in use.
BGP version	BGP version being used to communicate with the remote router. The neighbor's router ID (an IP address) is also specified.
BGP state	State of this BGP connection. (Refer to Chapter 15 for an explanation of the states.)
table version	Indication that the neighbor has been updated with this version of the primary BGP routing table.
up for	Amount of time that the underlying TCP connection has been in existence.
Last read	Time that the BGP process last read a message from this neighbor.
Hold time	Maximum amount of time that can elapse between messages from the peer or neighbor.
keepalive interval	Time period between sending keepalive packets, which maintain the TCP connection.
Received	Number of total BGP messages received from this peer, including keepalives.
Sent	Total number of BGP messages that have been sent to this peer, including keepalives.
notifications	Number of error messages that the router has sent to this peer.
Connections established	Number of times that the router has established a TCP connection for BGP between the two peers.
Dropped	Number of times that a valid TCP connection has failed or been taken down.
Connection state	State of BGP peer.
unread input bytes	Number of bytes of packets still to be processed.
Local host, Local port	Peering address of local router, plus its TCP port.
Foreign host, Foreign port	Neighbor's peering address and its TCP port.
Event Timers	Table that displays the number of starts and wakeups for each timer.
Iss	Initial send sequence number.

Table 16-3 Explanation of the **show ip bgp neighbors** Command (Continued)

Field	Descriptions
Snduna	Last send sequence number that the local host sent but for which an acknowledgment is outstanding.
Sndnxt	Sequence number that the local host will send next.
Sndwnd	TCP window size of the remote host.
Irs	Initial receive sequence number.
Rcvnxt	Last receive sequence number that the local host has acknowledged.
Rcvwnd	Local host's TCP window size.
delrcvwnd	Delayed receive window. Data that the local host has read from the connection but that has not yet been subtracted from the receive window that the host has advertised to the remote host. The value in this field gradually increases until it is larger than a full-sized packet, at which point it is applied to the rcvwnd field.
SRTT	A calculated smooth round-trip time.
RTTO	Round-trip time.
RTV	Variance of the round-trip time.
KRTT	New round-trip time (using the Karn algorithm). This field separately tracks the round-trip time of packets that have been retransmitted.
MinRTT	Smallest recorded round-trip time (hard-wire value used for calculation).
MaxRTT	Largest recorded round-trip time.
ACK hold	Time that the local host will delay an acknowledgment to piggyback data on it.
Flags	IP precedence of the BGP packets.
Datagrams	Header for the next two fields and the size of the Maximum Transmission Unit (MTU)
Rcvd	Number of update packets received from neighbor with data.
With Data Total Data Bytes	Total bytes of data.
Sent	Number of update packets sent.
With Data	Number of update packets with data sent.
Total data bytes	Total data bytes.

NOTE When implementing clusters and route reflectors, select the route reflector carefully in accordance with the physical topology of the network. Keep the design simple, placing one route reflector in a cluster. When the logical cluster design is in place, configure one cluster at a time and one route reflector at a time. After the route reflector in the cluster is configured, remove the BGP configuration that has the BGP sessions between the clients.

Controlling BGP Traffic

BGP updates can be controlled. It is often advantageous to limit the way that the BGP routing updates are propagated, for the same reasons that any routing protocol is best limited to only those updates that are required. This not only streamlines the traffic flow on the network, but also simplifies the network and thus its maintenance. Designing how the routing information should be forwarded through the network forms a basic level of security and can reduce the possibility of routing loops.

There are three main ways to apply policy-based routing in BGP:

- Making decisions based on the autonomous system path, the community, or the prefix
- Rejecting or accepting selected routes
- Setting attributes to influence the path selection

Rejecting or accepting selected routes requires some form of filtering through access lists. Filtering is used not only in policy-based routing, but also essentially as traffic control. There are three main flavors of filtering on a Cisco router:

- **Autonomous system path access list** —Used for filtering autonomous systems. An access list is used in BGP to filter updates sent from a peer based on the autonomous system path. In addition, other technologies use access lists for standard filtering.
- **Prefix list** —Used for filtering prefixes, particularly in redistribution. From Cisco IOS software version 11.2, ISPs were given prefix lists, which are a more efficient form of filtering. Prefix lists filter based on the prefix of the address. This option was made a part of IOS software version 12.0.
- **Distribute lists** —Used to filter routing updates. Although they are often used in redistribution, they are not specific to redistribution; they can be applied to inbound and outbound updates to or from any peer.

Both prefix lists and distribute lists filter on network numbers, not autonomous system paths, for which autonomous system path access lists are used.

- **Route maps** —Used to define routing policy. A route map is a sophisticated access list that defines criteria upon which a router acts when a match is found for the stated criteria. It is used in BGP for setting the attributes that determine the basis for selecting the best path to a destination.

Prefix lists are dealt with in more depth in the following sections. Route maps are dealt with in detail in Chapter 18, “Controlling Network Traffic with Route Maps and Policy-Based Routing,” and distribute lists are explained in Chapter 17, “Implementing Redistribution and Controlling Routing Updates.” Autonomous system path access lists are outside the scope of this book.

How Prefix Lists Work

Prefix lists were introduced in BGP because they are an efficient form of filtering. Because they search on the prefix of the address as defined by the administrator, the lookup is very fast. This is particularly important in the potentially huge routing tables that can be generated in BGP networks.

Another great advantage to prefix lists is the capability to edit them, particularly if they become large. Although it is possible to dynamically edit access lists, it is a little complicated. You must either port the access list to an application that allows editing or use named access lists.

Prefix lists are easier to create and use. This is true not only with the editing features, but also with the improved interface, which affords greater flexibility.

Before applying a prefix list to a process or an interface, you must first define the criteria for the prefix list. Each line in the prefix list is associated with a sequence number, similar to the number identifying a line of code in a computer program. If you choose not to enter the sequence number manually with the **prefix-list** command, the sequence numbers are automatically generated in increments of five. The sequence numbers that have not been used, for example, between 1 and 4, allow for additional lines to be added in subsequent edits of the prefix list. You can edit the prefix list by referencing the line or sequence number. This ability is not available in access lists, which require you to rewrite the entire list, unless you have the forethought to copy and paste the configuration file into a word processor.

Prefix lists work by matching the prefixes in the list to the prefixes of routes that are under scrutiny. The manner in which this is done is similar to that of access lists. When there is a match, the route is used or discarded.

More specifically, whether a prefix is permitted or denied is based upon the following rules:

- If a route is permitted, the route is used.
- If a route is denied, the route is not used.
- At the bottom of every prefix list is an implicit **deny any**. Thus, if a given prefix does not match any entries of a prefix list, it is denied.
- When multiple entries of a prefix list match a given prefix, the entry with the smallest sequence number (the first match in the list) is used.

- The router begins the search at the top of the prefix list, with the sequence number 1. When a match is made, the search stops. Processing time will be reduced if the most common matches or denies are placed near the top of the list. This will prevent having to process criteria that are seldom met every time a route is examined.
- Sequence numbers are automatically generated by default. To configure the sequence numbers manually, use the **seq seq-value** argument of the **ip prefix-list** command.
- A sequence number does not need to be specified when removing a configuration entry.

How to Configure a BGP Prefix List

Configuring a prefix list is straightforward if attention is given to the processing rules.

The following command creates an entry in a prefix list and assigns a sequence number to the entry:

```
Router(config-router)#ip prefix-list prefix-list-name [seq seq-value ] {deny | permit}
network/len [ge ge-value] [le le-value]
```

Table 16-4 explains the parameters shown in the preceding syntax.

Table 16-4 Explanation of the ip prefix-list Command

Parameter	Description
<i>prefix-list-name</i>	Name of the prefix list
[<i>seq seq-value</i>]	The sequence number to be assigned to the criteria being defined
{ deny permit }	Whether the action is to deny or permit the route based on the outcome of the criteria
<i>network/len</i>	The length of the prefix length to be matched
[ge ge-value] [le le-value]	The optional keywords ge and le can be used to specify the range of the prefix length to be matched. An exact match is assumed when neither ge nor le is specified

To configure a router to use a prefix list as a filter in distributing routes, use the following command:

```
Router(config-router)#neighbor {ip address | peer-group} prefix-list prefix-list-name
{in | out}
```

An example of a simple prefix list follows:

```
Router(config)#ip prefix-list tryout permit 44.0.0.0/8
Router(config)#ip prefix-list tryout permit 130.0.0.0/8
```

The prefix list “tryout” will allow the networks 44.0.0.0 and the supernet 130.0.0.0 to be further processed by BGP.

Sometimes it is necessary to create a criteria range as opposed to an absolute. For example, you could change “all 2-year-old children are allowed into the playground” to “children between the ages of 2 and 4 are allowed into the playground.” This grants greater flexibility to the searches. The way to do this in a prefix list is to use the **ge** and **le** parameters.

These optional keywords allow a range of the prefix length to be specified, as opposed to the *network/len*, which is the absolute. Therefore, 10.2.3.0/24 is an example of the *network/len*, which states the prefix to be matched and the length of the prefix. The equations are confusing until you sit and work them out. The following are some key points:

- **ge** is used if the prefix is greater than or equal to the value stated in the list.
- **le** is used if the prefix is less than or equal to the value stated in the list.

Simply put, the *ge-value* is the barrier for the lower limit, in that the number must be greater than the value stated in for the *ge-value*. Likewise, the *le-value* is the barrier for the upper limit, in that the number must be less than that stated in the *le-value*. So, children entering the playground must be older than 2 (*ge-value* of 2) and younger than 4 (*le-value* of 4). Therefore, the formula requires the following condition:

$$\text{len} < \text{ge-value} < \text{or} = \text{le-value} < \text{or} = 32$$

For example, to permit all prefixes between /8 and /24, you would use the following:

```
Router(config)#ip prefix-list tryone permit 0.0.0.0/0 ge 8 le 24
```

NOTE An exact match is assumed when neither **ge** nor **le** is specified. The range is assumed to be from *ge-value* to 32 if only the **ge** attribute is specified, and from *len* to *le-value* if only the **le** attribute is specified.

Verifying the Prefix List Configuration

As always, it is important to check the configuration, especially if it involves the filtering of routes or routing updates. Table 16-5 lists the various **show** commands available for prefix lists.

To display information about a prefix list or prefix list entries, use the **show ip prefix-list** exec command:

```
Router#show ip prefix-list [detail|summary] name [network/len] [seq seq-num]
[longer] [first-match]
```

Table 16-5 *Displaying Prefix List Command Options*

Command	Description
show ip prefix-list [detail summary]	Displays information about all prefix lists, including the hit count, which is the number of times that a match has been found for the criteria in the prefix list. This is very important in troubleshooting for capacity planning and security.
show ip prefix-list [detail summary] <i>name</i>	Displays a table showing the entries in a prefix list identified by name.
show ip prefix-list <i>name</i> [network/len]	Displays the filtering associated with the node based on the absolute of the defined prefix.
show ip prefix-list <i>name</i> [seq seq-num]	Displays the prefix list entry with a given sequence number.
show ip prefix-list <i>name</i> [network/len] longer	Displays all entries of a prefix list that are more specific than the given network and length.
show ip prefix-list <i>name</i> [network/len] first-match	Displays the entry of a prefix list that matches the given prefix (network and length of prefix).

The **show** commands always include the sequence numbers in their output.

Example 16-4 shows the output of the **show ip prefix-list** command with details about the prefix list “tryout.”

Example 16-4 *A Sample Output of the show ip prefix-list Command*

Router# show ip prefix-list detail tryout Ip prefix-list tryout: Description: tryout-list count: 1, range entries: 0, sequences: 5 - 10, refcount: 3 seq 5 permit 44.0.0.0/8 (hit count: 0, refcount: 1) seq 10 permit 130.0.0.0/8 (hit count: 0, refcount: 1)

As you have seen, filtering routes is crucial in the design of BGP in order to maintain manageable routing tables and to conserve network resources. Up to this point, the discussion has centered on the use of BGP within and between autonomous systems. When connecting to the Internet, you need to consider slightly different design issues, as discussed in the next section.

Connecting to the Internet with BGP

As an exterior routing protocol, BGP is used to connect to the Internet and to route traffic within the Internet. You need to be aware of some design considerations when connecting to such an enormous resource. Just like driving a car during rush hour, you need different skills depending on whether you are joining a freeway or driving through a small town. When connecting to the Internet, your network is joining a major freeway that connects large cities, so you must ensure that the amount of traffic and information does not overwhelm your network.

The following sections deal with two important design considerations: the need for redundant links into the Internet, called multihoming, and the need to decide how much information to receive from the Internet.

Redundant Connections to the Internet—Multihoming

An enormous amount of traffic leaves an organization in search of Internet resources. This traffic includes not only e-mail and other means of communication, but also requests for information from the Internet.

Use of the Internet continues to expand as both an individual tool and a major mechanism of finance and commerce. It becomes increasingly necessary for the network administrator to provide constant access to the Internet with load balancing and redundancy with multihoming.

To have more than one connection to the Internet is to be *multihomed*. The reason for duplicating the connection is clear: The need for Internet access nowadays is too great for the responsibility to fall onto one link. Multiple links not only provide redundancy, but also allow for load balancing and thus present an improvement in performance.

Multihoming might be several connections to the same ISP, or it might include another layer of redundancy by making the second connection to another ISP. The following are some concerns about connecting to more than one ISP:

- Each provider might not be propagating the same routes into or from the Internet. If the providers are sending subsets of the required routes, there could be a major problem with connectivity if the link to one of the providers fails.
- If you are connected to two different providers, your autonomous system could become a transit autonomous system between the ISPs. This could happen if a router in the autonomous system of one provider sees a path to a destination via the other provider's autonomous system, and your autonomous system gives the best route to the autonomous system of the other provider.

Configuration at the ISP level is the solution to these concerns and is dealt with when setting up the service. Therefore, it is important that you raise the need for multihoming during negotiations with the ISP so that the ISP is aware of the need for additional configuration.

Receiving Routing Information from the Internet

When connecting to something as vast as the Internet, some planning and forethought is necessary. In particular, it is essential to decide what updates are to be sent to the outside world and how routers within the autonomous system are to know about the outside world and all that it offers.

There are three main approaches to the selection of routes from the Internet:

- Accept only default routes from all providers
- Accept partial routes in addition to default routes from all providers
- Accept full routing updates from all providers

The decision process is clear: It is a balance of network resources against information. The greater the amount of resources, the more routes can be accepted from the providers.

Table 16-6 summarizes the different approaches to obtaining routing information from the Internet.

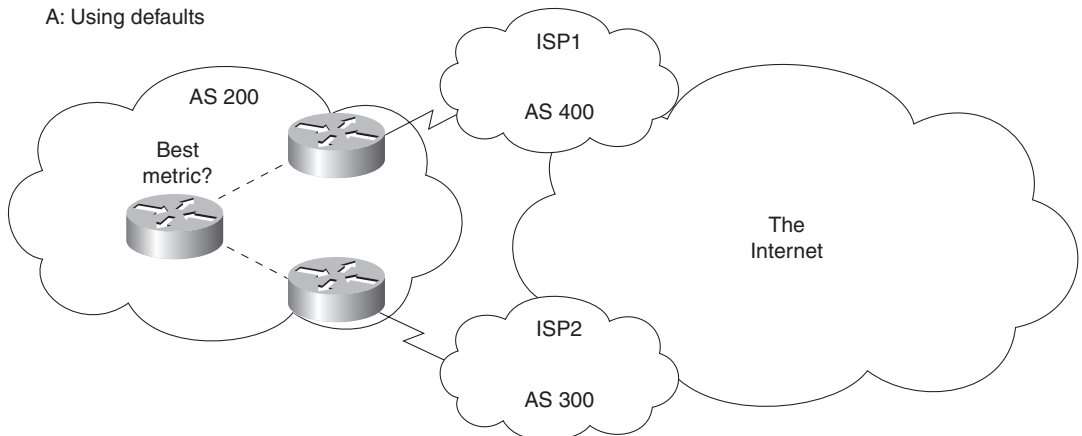
Table 16-6 Receiving Routing Updates from Multiple ISPs

Routes Accepted from the Internet	Memory	CPU	IGP Chooses Best Metric to Default Network?	BGP Selects Best Path to External Network?	Can Exit Path Be Tuned via BGP Attributes?	AS Sends All Its Routes to ISPs?	ISP Chooses Entry Path to AS?
Default routes only from all ISPs	Low	Low	Yes	Go to nearest gateway that is advertising the route	No	Yes	Yes
Select routes and default routes from ISPs	Medium	Medium	Yes	Yes; normally the AS_Path is the attribute that selects the exit path to the Internet	Yes	Yes	Yes
Full routing tables from ISPs	High	High	Yes	Yes; normally the AS_Path is the attribute that selects the exit path to the Internet	Yes	Yes	Yes

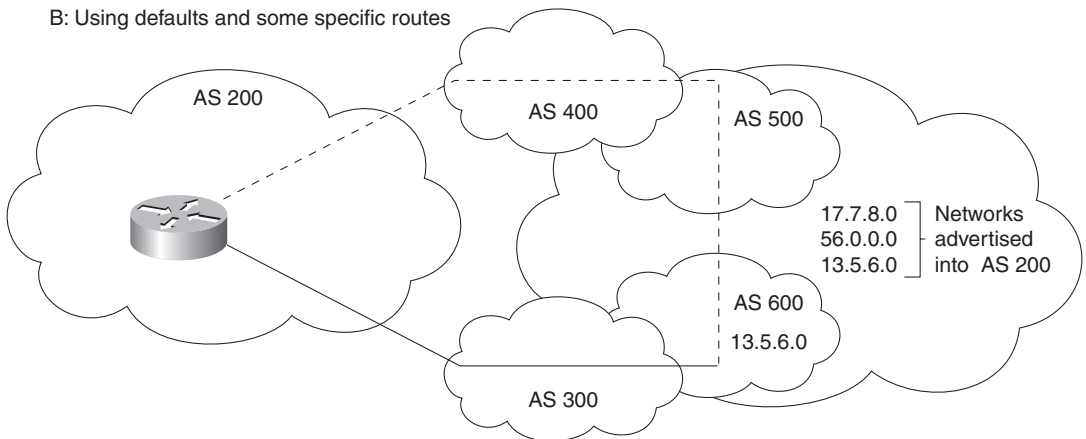
Figure 16-7 illustrates the various options available in exchanging routing information with the Internet.

Figure 16-7 *Exchanging Routing Information with the Internet*

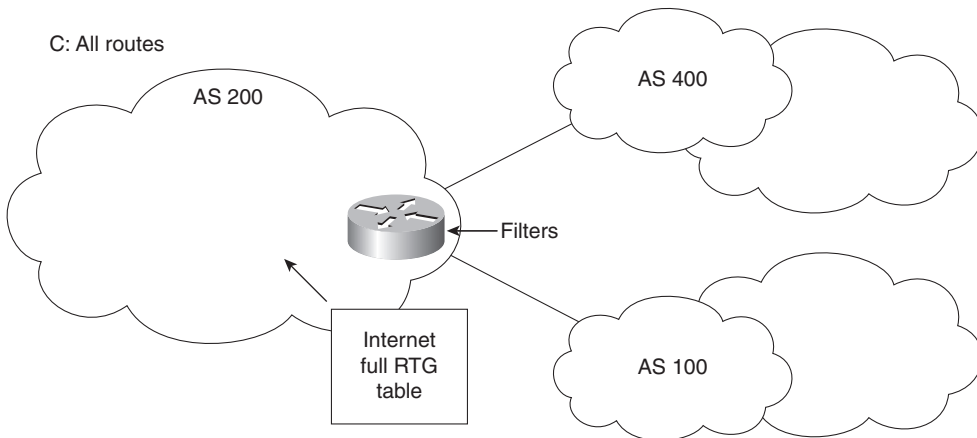
A: Using defaults



B: Using defaults and some specific routes



C: All routes



NOTE The second solution, accepting partial routes from the ISP, requires the updates sent into the autonomous system to be filtered, either by your autonomous system or by the ISP. If the responsibility falls to your organization, you will need to study the use of route maps and regular expressions. This is a complex subject, which is explained in detail on the Cisco web site. As of press time for this book, the best information on this subject could be found at <http://www.cisco.com/warp/public/459/27.html>. Or, go to Cisco.com and search for “sample configurations for BGP.”

Determining the BGP Path by Tuning the Attributes

Now that you understand the network requirements for designing and configuring a BGP network, as well as methods for controlling BGP traffic through the network, you will learn how to configure BGP to take a path to a destination based on different criteria.

The attributes discussed in this section are local preference and weight, with the latter being a Cisco-proprietary solution.

Using the Local Preference and Weight Attributes

The weight attribute selects the exit path out of the router when there are multiple paths to the same destination. The higher the weight value, the better the path. This command is a local command, and the attribute is not propagated to other routers. It is also proprietary to Cisco. To configure the weight attribute, use the following command:

```
Router(config-router)#neighbor {ip-address | peer-group-name} weight weight
```

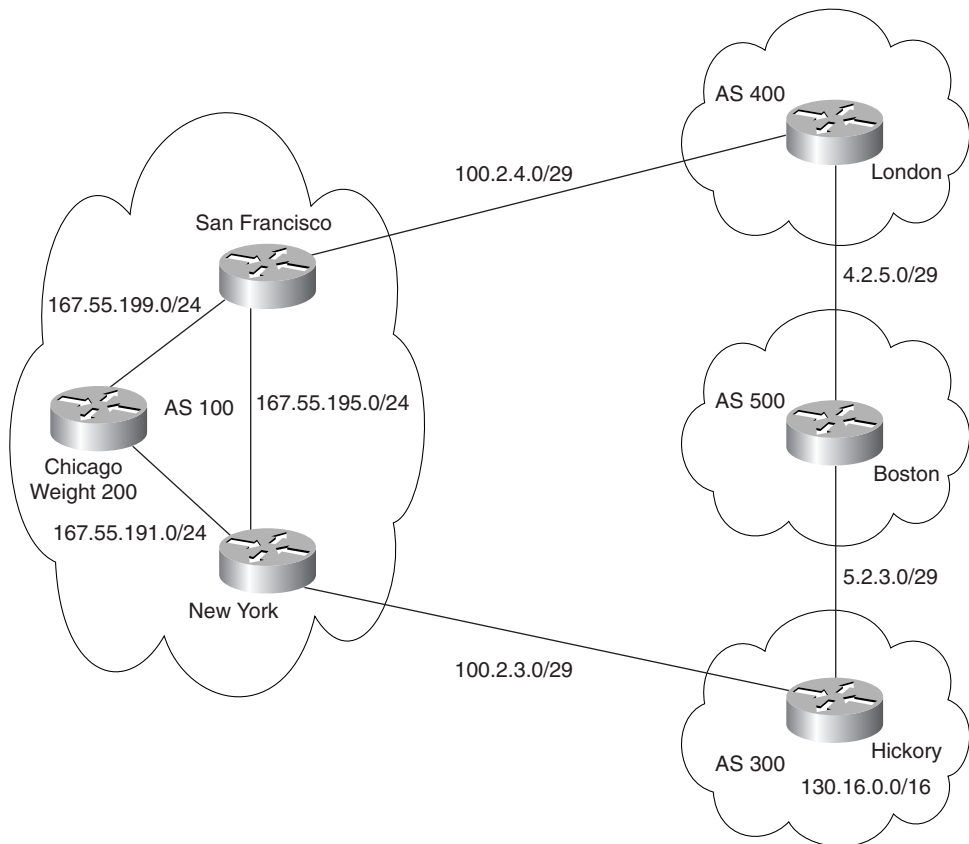
Table 16-7 explains the meaning of the preceding syntax.

Table 16-7 *An Explanation of the Command to Configure the Weight Attribute*

Syntax	Description
neighbor	This indicates that the rest of the command is directed at a BGP peer.
<i>ip-address</i>	This is the IP address of the neighboring router.
<i>peer-group-name</i>	This identifies the BGP peer group, if there is one.
weight weight	This is proprietary to Cisco and is used in route selection. It is local to the router, and because it is not propagated to other routers, there is no problem with compatibility. When there are multiple paths, it selects a path to a destination with different next hops to the same destination. This identifies the weight attribute, and a value is placed immediately afterward. The default is 32,768, although the range extends from 0 to 65,535, and higher is preferred over lower.

Figure 16-8 illustrates the use of the weight attribute, and Example 16-5 shows how the path through San Francisco is chosen. As you can see, the weight has been set on Chicago, making it prefer the path through San Francisco no matter which network it is trying to reach. The best path to 130.16.0.0 is through New York. (Refer to Example 16-7 later in the chapter to see that the **show ip bgp** command on Chicago shows that the weight attribute forces San Francisco to be used as the next hop for all paths.)

Figure 16-8 *The Weight Attribute and Selecting a Path*



Example 16-5 *A Sample Configuration to Illustrate How to Tune the Weight Attribute*

```
!
Chicago(config)#router bgp 100
!
Chicago(config-router)#bgp log-neighbor-changes
Chicago(config-router)#neighbor 167.55.191.3 remote-as 100
```

continues

Example 16-5 *A Sample Configuration to Illustrate How to Tune the Weight Attribute (Continued)*

```
Chicago(config-router)#neighbor 167.55.199.2 remote-as 100
Chicago(config-router)#neighbor 167.55.199.2 weight 200
!The route from Chicago to San Francisco will have a weight of 200, making this the path that
  is chosen
Chicago#
```

The local preference is equally easy to configure. You can set it on either a default or a per-prefix basis. The command to set the local preference on per-prefix basis follows:

```
Router(config-route-map)# set local-preference local-preference
```

The syntax for the default command is as follows:

```
Router(config-router)#bgp default local-preference value
```

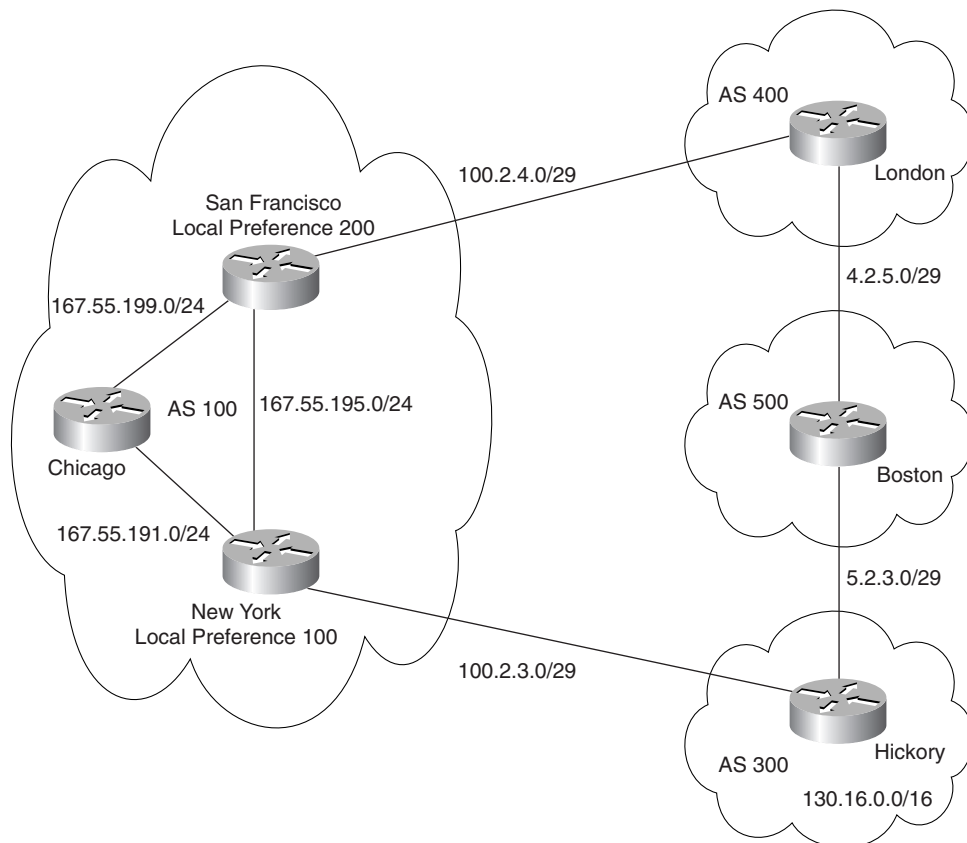
Table 16-8 explains the various parts of the default configuration command. The command to set the local preference on per-prefix basis was provided for reference only because it is outside the scope of this book.

Table 16-8 *Configuring the Local Preference Attribute*

Syntax	Description
bgp default local-preference	This attribute is used to tell routers within the autonomous system how to exit the autonomous system in the case of multiple paths. It is the opposite of the MED attribute.
<i>value</i>	Local preference has a range from 0 to 4,294,967,295 (just over 4 billion). The higher value is preferred in selecting routes. The default is 100.

Example 16-6 is based on Figure 16-9. The local preference, set in the San Francisco router to 200, is propagated in the updates to all its peers. Likewise, the local preference of 100 set in the New York router is propagated to its peers. When Chicago has to decide on a path to the network 130.16.0.0, the highest local preference attribute dictates the San Francisco router as the exit point from the autonomous system.

The configurations in Examples 16-5 and 16-6 are extremely simple. Although they work well, you also need to understand route maps. Route maps allow the setting of attributes with conditions and other criteria. Although they are more complex, they are also more efficient. Route maps offer more flexibility with their greater level of sophistication. Route maps are explained in detail in Chapter 18.

Figure 16-9 *Using Local Preference to Select a Path***Example 16-6** *A Sample Configuration to Illustrate How to Tune the Local Preference Attribute*

```

SanFrancisco(config)#router bgp 100
!
SanFrancisco(config-router)#bgp default local-preference 200
SanFrancisco(config-router)#aggregate-address 167.55.0.0 255.255.0.0 summary-only
SanFrancisco(config-router)#neighbor 100.2.4.4 remote-as 400
SanFrancisco(config-router)#neighbor 100.2.4.4 default-originate
!
SanFrancisco(config-router)#neighbor 167.55.195.3 remote-as 100
SanFrancisco(config-router)#neighbor 167.55.199.1 remote-as 100
!
NewYork(config)#router bgp 100
!

```

continues

Example 16-6 *A Sample Configuration to Illustrate How to Tune the Local Preference Attribute (Continued)*

```
NewYork(config-router)#bgp default local-preference 100
NewYork(config-router)#network 167.55.0.0
NewYork(config-router)#aggregate-address 167.55.0.0 255.255.0.0 summary-only
NewYork(config-router)#neighbor 100.2.3.2 remote-as 300
NewYork(config-router)#neighbor 167.55.191.1 remote-as 100
NewYork(config-router)#neighbor 167.55.195.2 remote-as 100
!
```

Verifying the Configuration of Attributes

It is always important to be able to check your work, particularly when that work defines an entire organization’s method of connecting to the Internet.

The **show ip bgp** command shows all the values of all the attributes and their status. Therefore, this is a good command to verify any configurations that change attributes to tune the system and effectively manage the traffic flow to and from the autonomous system.

Examples 16-7 through 16-9 show sample output from the **show ip bgp** command.

Example 16-7 shows how BGP is running before the configuration in Example 16-5 or Example 16-6 has been run on the Chicago or San Francisco router. The next hop is to 100.2.3.2, which is in autonomous system 300 because the traffic would be routed via New York. Note in Example 16-7 that the local preference on Chicago has been set by the BGP process to be 100 by default.

Example 16-7 *The show ip bgp Command Example for Chicago Before Attributes Are Set*

```
Chicago#show ip bgp
BGP table version is 22, local router ID is 192.168.0.231
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop          Metric LocPrf Weight Path
*>i4.0.0.0          100.2.4.4              0    100      0 400 I
*>i5.0.0.0          100.2.3.2              0    100      0 300 I
*>i100.2.3.0/29     100.2.3.2              0    100      0 300 I
*>i100.2.4.0/29     100.2.4.4              0    100      0 400 I
*>i130.16.0.0       100.2.3.2              0    100      0 300 I
r>i167.55.0.0       167.55.191.3          281600  100      0 I
Chicago#
```

Example 16-8 occurs after the configuration for Chicago; the weight is set to 200 for the neighbor 167.55.199.5, which is San Francisco. This forces the longest path to be taken to 130.16.0.0 via San Francisco.

Example 16-8 *The show ip bgp Command Example Showing the Use of the Weight Attribute*

```
Chicago#show ip bgp
BGP table version is 8, local router ID is 192.168.0.231
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop           Metric LocPrf Weight Path
*>i4.0.0.0          100.2.4.4              0    100    200 400 i
*>i5.0.0.0          100.2.3.2              0    100      0 300 i
*>i100.2.3.0/29     100.2.3.2              0    100      0 300 i
*>i100.2.4.0/29     100.2.4.4              0    100    200 400 i
*>i130.16.0.0       100.2.4.4             100    200 400 500 300 i
* i                 100.2.3.2              0    100      0 300 i
r>i167.55.0.0       167.55.191.3          281600  100      0 i
Chicago#
```

Example 16-9 occurs after the San Francisco and New York routers are configured; it is possible to see a change in the neighbor table. The local preference shows that BGP packets destined to 130.16.0.0 still take the high road, against common sense, because the local preference instructs BGP that San Francisco has the best path. This attribute is propagated to other BGP neighbors. You can see that there is only one path shown in this table to 130.16.0.0. Because both New York and San Francisco agree that the path through New York is inferior, it is not sent. Note that the local preference to 167.55.191.3, known to its friends as New York, has a local preference of 100.

Example 16-9 *The show ip bgp Command Example Showing the Use of the Local Preference Attribute*

```
Chicago#show ip bgp
BGP table version is 24, local router ID is 192.168.0.231
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop           Metric LocPrf Weight Path
*>i4.0.0.0          100.2.4.4              0    200      0 400 i
*>i5.0.0.0          100.2.4.4             200      0 400 500 300 i
r>i100.2.3.0/29     100.2.4.4             200      0 400 500 300 i
*>i100.2.4.0/29     100.2.4.4              0    200      0 400 i
*>i130.16.0.0       100.2.4.4             200      0 400 500 300 i
r>i167.55.0.0       167.55.191.3          281600  100      0 i
Chicago#
```

Table 16-9 describes significant fields shown in Examples 16-7 through 16-9.

Table 16-9 *Explanation of Output from the `show ip bgp` Command*

Field	Description
BGP table version	Internal version number of the table. This number is incremented whenever the table changes.
Local router ID	The highest IP address of the router.
Status codes	Status of the table entry. The status is displayed at the beginning of each line in the table. It can be one of the following values: s—The table entry is suppressed. *—The table entry is valid. >—The table entry is the best entry to use for that network. i—The table entry was learned via an iBGP session. D—The table entry was dampened.
Origin	The origin of the entry. The origin code is placed at the end of each line in the table. It can be one of the following values: i—Entry originated from IGP and was advertised with a network router configuration command. e—Entry originated from an EGP. ?—Origin of the path is not clear. Usually, this is a router that is redistributed into BGP from an IGP.
Network	A destination IP address of a network.
Next hop	IP address of the next logical device in the path to the destination. The forwarded packets are sent to this address. An entry of 0.0.0.0 indicates that the router has some non-BGP routes to this network.
Metric	If shown, this is the value of the metric between autonomous systems. This field is frequently not used.
LocPrf	Local preference values as set on the routers with interfaces to other autonomous systems. It defines how preferable that router is as a transit point out of the autonomous system. The default value is 100.
Weight	Weight of the route, determining which path the router will choose. It is proprietary to Cisco and is an attribute local to the router.
Path	Autonomous system paths to the destination network. There can be one entry in this field for each autonomous system in the path.

Redistribution Between IGP and BGP

Up until now, the discussion has been about BGP, its operation, and the configuration options available. However, for most networks, BGP is the means by which information about the Internet is brought into the internal organizational routing domain. Disseminating this information throughout the autonomous system is the subject of the next section.

If the organization is not an ISP, there is a fair chance that the network is running an IGP within the autonomous system. The IP routing table generated by this protocol or protocols is distinct from the BGP routing table, although as you have seen, they communicate freely. The use of synchronization is a case in point. However, populating one routing table with routes from the other routing table needs to be manually configured.

Routes can be injected from the IGP into BGP, or from the BGP into the IGP, as discussed in the next sections.

Advertising Routes from IGP into BGP

First, consider the advertising routes into BGP. There are three ways of populating the BGP table with IGP routes:

- **Using the network command** — This is used to advertise routes that are in the IP routing table.
- **Redistributing static routes** — Although any static route can be redistributed into BGP, static routes are often used to create a supernet. The static route is a summary of classful addresses, such as Class C addresses with a prefix mask of 16 bits instead of 24 bits. This requires statically routing to null 0. This fools the system by creating a route that has no exit point from the router because the route does not exist but is redistributed into BGP. The command places the route into the routing table without fear of it being used and creating a black hole.

The problem is that if the route in the IGP routing table disappears, BGP still advertises the route, causing traffic to journey into the autonomous system, only to die. Therefore, Cisco suggests that you use the **aggregate-address** command for BGP instead.

- **Redistributing dynamically learned routes from the IGP** — This configuration is not advised because there is a great reliance on the IGP table. It is imperative that external routes carried in iBGP are filtered out; otherwise, routing loops are generated when BGP routes are fed into IGP, only to be advertised back into BGP further down the network.

Advertising Routes from BGP into an IGP

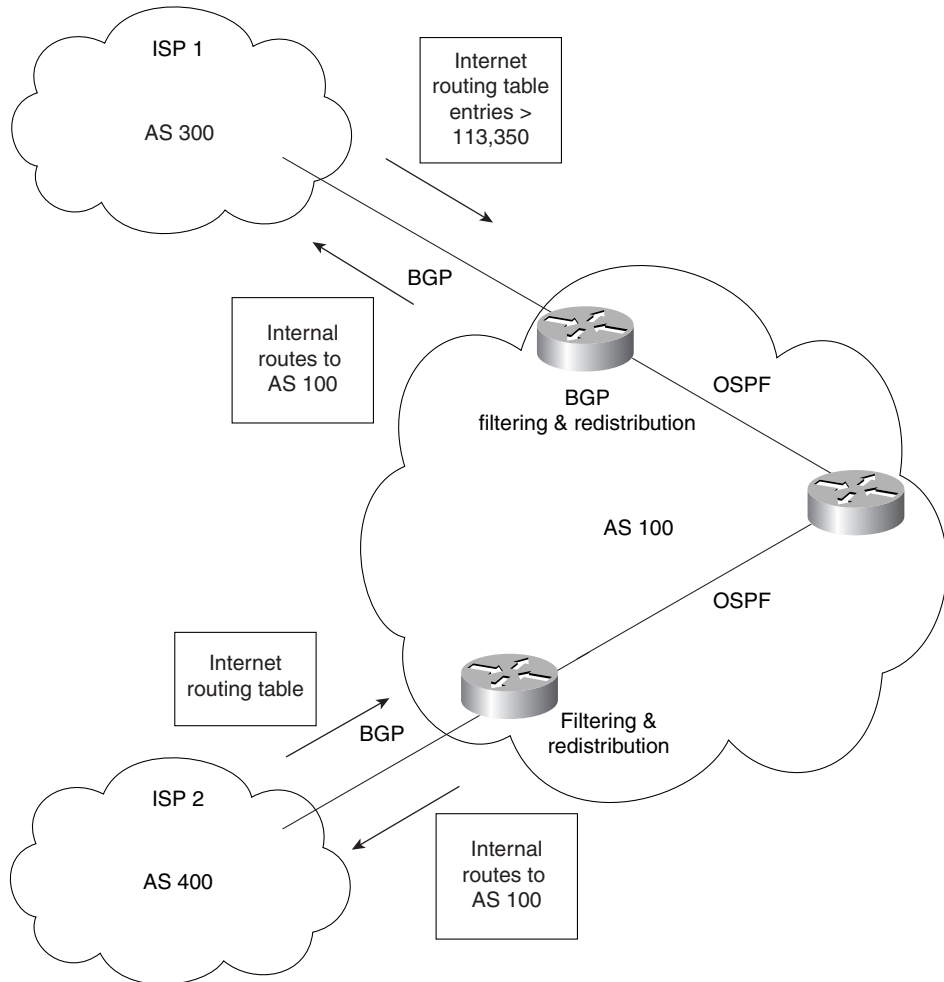
Redistributing the routes from the Internet into a small network is unwise. The proposition is alarming because of the vastness of the Internet and the enormity of the routing tables. Even with a large amount of aggregation and filtering, there is still a lot of information to carry over.

Because ISPs tend to run eBGP and iBGP extensively, ISPs often run them exclusively for exterior routes, using an IGP only for internal routes. This requires no redistribution, which is easier for the routers and means the following:

- The resources are available for other processes.
- The IGP routing table is spared.

The synchronization function is not necessary in this type of network because iBGP is running in a fully meshed environment. With either a fully meshed network or carefully designed route reflectors, the synchronization rule can be turned off.

If iBGP is multihomed, redistribution from BGP into the IGP is needed. The IGP needs to carry the external routes across the autonomous system to the other BGP router. Also, any device wanting to connect to the Internet needs to have either a default route or specific routes to direct traffic forward. Filtering must be configured; otherwise, the internal routing tables will become overwhelmed. This is illustrated in Figure 16-10.

Figure 16-10 *Redistributing BGP Routes into a Non-ISP Organization*

Foundation Summary

The “Foundation Summary” section of each chapter lists the most important facts from the chapter. Although this section does not list every fact from the chapter that will be on your exam, a well-prepared candidate should, at a minimum, know all the details in each “Foundation Summary” before going to take the exam.

Table 16-10 summarizes the commands you have seen throughout this chapter.

Table 16-10 *Summary of Commands Used in This Chapter*

Command	Description
neighbor <i>ip-address</i> route-reflector-client	Configures the route reflector to view a specified neighbor as a client. It reflects all routes received from this router to other clients and nonclients.
no neighbor <i>ip-address</i> route-reflector-client	Removes a neighbor as a client.
show ip bgp neighbor <i>neighbor-address</i>	Gives details about a specified neighbor.
ip prefix-list <i>list-name</i> [seq <i>seq-value</i>] deny permit <i>network/len</i> [ge <i>ge-value</i>] [le <i>le-value</i>]	Creates a prefix list that is used to filter routes from updates based on the prefix length of the route.
neighbor <i>ip-address</i> <i>peer-group-name</i> prefix-list <i>prefix-list-name</i> in out	Configures a router to use a prefix list as a filter in distributing routes.
show ip prefix-list detail <i>name</i>	Shows a prefix list by name. It shows what the list is filtering on and how many matches to the criteria have been logged. This is known as the number of hits.
neighbor { <i>ip-address</i> <i>peer-group-name</i> } weight <i>weight</i>	Sets the weight attribute to influence BGP path selection. Weight is a Cisco-proprietary attribute. It is a local attribute, which is not propagated to other routers. It selects the best path if there are multiple exits from the router into the Internet.
bgp default local-preference <i>value</i>	Sets the local preference attribute. Again, this is used to influence the router’s selection of the best path based on its selection process. This attribute is passed in the routing updates to other BGP peers.
show ip bgp	Shows details about the BGP configuration on your network.

The benefits of route reflectors include the following:

- The capability to scale the network, given the other characteristics
- A strong hierarchical design
- A reduction of traffic on the network
- A reduction in the memory and CPU needed to maintain TCP sessions
- Faster convergence and a simpler network because two routing protocols are implemented:
 - iBGP for external routing information traversing the autonomous system
 - IGP for routes internal to the autonomous system

Characteristics of route reflectors are as follows:

- A route reflector is a router that forwards updates to its clients. When a client sends an update to the route reflector, it is forwarded or reflected to the other clients and nonclients.
- The route reflector is the only router that is configured or that has the remotest idea that it is anything other than a peer.
- A client is a router that receives updates from a route reflector that a route reflector has forwarded from another client or nonclient.
- Both a route reflector and its clients, therefore, form a unit that shares information. This unit is called a cluster.
- The autonomous system can be divided into clusters and be configured. There must be at least one route reflector per cluster; otherwise, the clients will not get the updates reflected to them.
- The route reflector and the client no longer require a full mesh of peering relationships because the route reflector forwards updates from other clients.
- In all probability, a route reflector is connected to peers for whom it is not forwarding routes. These are regular neighbors or peers, but from the route reflector's view, they are nonclients.
- Nonclients must be fully meshed with the route reflector and each other.
- The route reflector connects to other route reflectors. These route reflectors need to be fully meshed because the old rule of not propagating routes that are not defined in the **network** command is now operational. This is to ensure that the iBGP routing tables are complete.
- When the route reflector forwards an update, the Originator-ID attribute is set. This is the BGP router ID of the router that originated the path. The purpose of this attribute is not to award honors to the originating router, but so that if this router receives the update, it will see its own ID and will ignore the packet. This prevents the possibility of routing loops.
- If there are multiple route reflectors in the cluster to provide redundancy, then the originating router is identified by the Cluster-ID attribute. This serves the same purpose as the Originator-ID in preventing routing loops.

The rules by which route reflectors propagate updates are as follows:

- If a route reflector receives multiple paths to the same destination, it chooses the best path.
- If the route is received from a client, the route reflector reflects or forwards the update to clients and nonclients, except for the originator of the route.
- If the route is received from a nonclient, the route reflector reflects the update only to clients.
- If the route is received from eBGP, the route reflector or client reflects it to all nonclients, as well as clients.

Whether a prefix is permitted or denied is based upon the following rules:

- If a route is permitted, the route is used.
- If a route is denied, the route is not used.
- At the bottom of every prefix list is an implicit **deny any**. Thus, if the given prefix does not match any entries of a prefix list, it is denied.
- When multiple entries of a prefix list match a given prefix, the entry with the smallest sequence number (the first match in the list) is used.
- The router begins the search at the top of the prefix list, with the sequence number 1. When a match is made, the search stops. Processing time will be reduced if the most common matches or denies are placed near the top of the list. This will prevent having to process criteria that are seldom met every time a route is examined.
- Sequence numbers are generated automatically by default. To configure the sequence numbers manually, use the **seq seq-value** argument of the **ip prefix-list** command.
- A sequence number does not need to be specified when removing a configuration entry.

Table 16-11 lists the various command options for prefix lists.

Table 16-11 *Displaying Prefix List Command Options*

Command	Description
show ip prefix-list [detail summary]	Displays information about all prefix lists, including the hit count, which is the number of times that a match has been found for the criteria in the prefix list. This is very important in troubleshooting for capacity planning and security.
show ip prefix-list [detail summary] name	Displays a table showing the entries in a prefix list identified by name.
show ip prefix-list name [network/len]	Displays the filtering associated with the node based on the absolute of the defined prefix.

Table 16-11 *Displaying Prefix List Command Options (Continued)*

Command	Description
show ip prefix-list <i>name</i> [seq <i>seq-num</i>]	Displays the prefix list entry with a given sequence number.
show ip prefix-list <i>name</i> [<i>network/len</i>] longer	Displays all entries of a prefix list that are more specific than the given network and length.
show ip prefix-list <i>name</i> [<i>network/len</i>] first-match	Displays the entry of a prefix list that matches the given prefix (network and length of prefix).

The **show** commands always include the sequence numbers in their output.

Table 16-12 summarizes the different approaches to obtaining routing information from the Internet.

Table 16-12 *Receiving Routing Updates from Multiple ISPs*

Which Routes Are Accepted from the Internet	Memory	CPU	IGP Chooses Best Metric to Default Network?	BGP Attribute Selects Best Path to External Network?	Can Exit Path Be Tuned via BGP Attributes?	Autonomous System Sends All Its Routes to ISPs?	ISP Chooses Entry Path to Autonomous System?
Default routes only from all ISPs	Low	Low	Yes	Go to nearest gateway that is advertising the route	No	Yes	Yes
Select routes and default routes from ISPs	Medium	Medium	Yes	Yes; normally the AS_Path is the attribute that selects the exit path to the Internet	Yes	Yes	Yes
Full routing tables from ISPs	High	High	Yes	Yes; normally the AS_Path is the attribute that selects the exit path to the Internet	Yes	Yes	Yes

Q&A

As mentioned in the introduction, “All About the CCNP, CCDP, and CCIP Certifications,” you have two choices for review questions. The questions that follow next give you a bigger challenge than the exam itself by using an open-ended question format. By reviewing now with this more difficult question format, you can exercise your memory better and prove your conceptual and factual knowledge of this chapter. The answers to these questions are found in Appendix A.

For more practice with examlike question formats, including questions using a router simulator and multichoice questions, use the exam engine on the CD-ROM.

1. If a route reflector hears an update from a nonclient, what action will be taken?
2. In version 11.0 of the Cisco IOS software, what method would be used to restrict routing information from being received or propagated?
3. Explain the purpose and use of the command **show ip prefix-list** *name [seq seq-number]*.
4. Why would you redistribute static routes into BGP?
5. Why is it advisable to have the route reflectors fully meshed?
6. Why is filtering often required when redistributing BGP into an IGP?
7. What are the advantages of multihoming?
8. Why do iBGP peers need to be fully meshed?
9. How is a fully meshed network avoided in iBGP?
10. What is the equation to determine the number of sessions needed in a fully meshed BGP network?
11. Why does a fully meshed network in iBGP cause problems?
12. State two benefits to using route reflectors.
13. If a route reflector sees multiple paths to a destination, what action is taken?
14. Explain the difference between a cluster-ID and an originator-ID.
15. State two advantages in using prefix lists over access lists.
16. If the ISP has provided a default route, how will the router within the autonomous system select the exit path in a multihomed environment?
17. What is a disadvantage of an autonomous system receiving full routing updates from all ISPs?
18. What is the danger of redistributing BGP into the IGP?

19. What are the advantages of a fully meshed iBGP network?
20. In configuring a route reflector, how is the client configured?
21. What commands are used to display the BGP router ID that identifies the router that is sending the updates and peering with its neighbor?

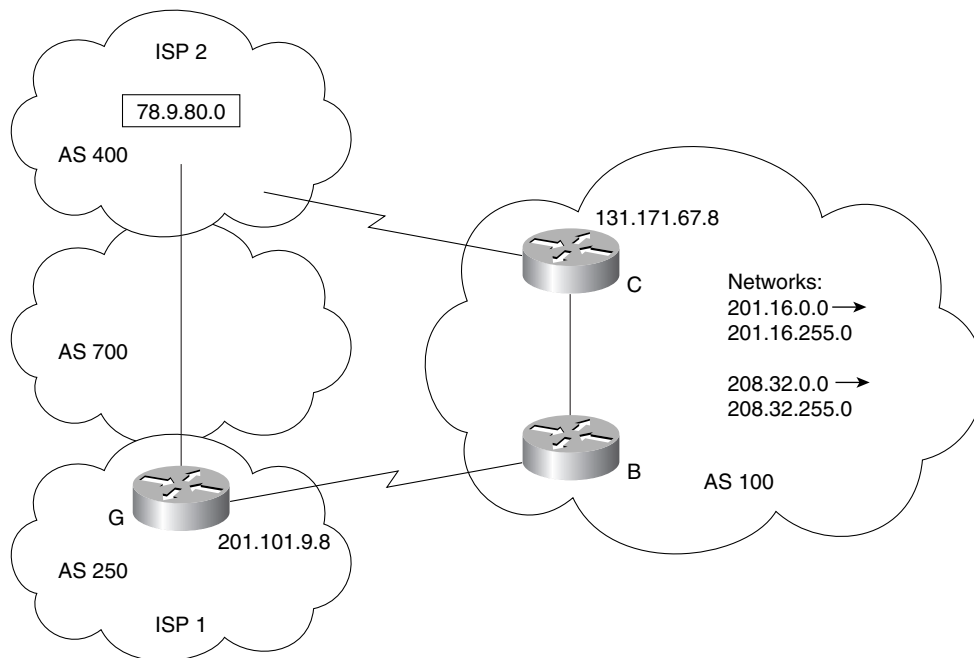
Scenarios

The following scenarios and questions are designed to draw together the content of the chapter and to exercise your understanding of the concepts. There is not necessarily a right answer. The thought process and practice in manipulating the concepts are the goals of this section. The answers to the scenario questions are found at the end of this chapter.

Scenario 16-1

The company Humugos has successfully implemented iBGP in each country that it operates in, with eBGP connecting the autonomous systems. The company now wants to change the way it connects to the Internet. Currently, it has one connection into the Internet per autonomous system. Figure 16-11 provides the diagram for the network in this scenario.

Figure 16-11 *Diagram for Scenario 16-1*



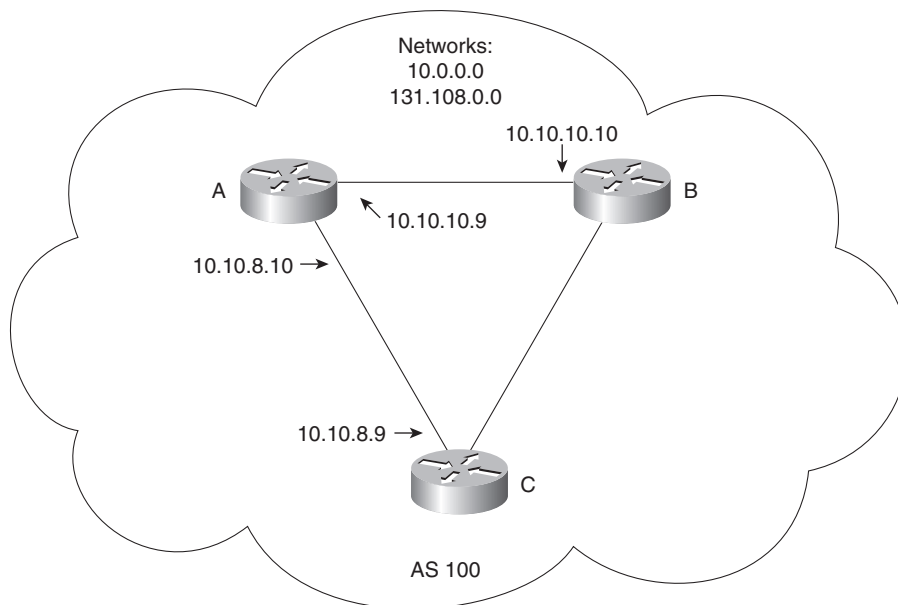
1. Give reasons to support Humugos's desire to have multiple connections to the Internet.
2. The company has been advised to redistribute static routes into the Internet BGP. It had intended to redistribute dynamic OSPF routes directly into the ISP provider. Explain why the ISP was not in favor of this configuration.
3. Using Figure 16-11, issue the configuration commands that would allow Router B connecting into the Internet to select the path to network 78.9.80.0 via Router G. Use the local preference attribute to select the path.

NOTE This network scenario is oversimplified for learning purposes. Normally, it would be very difficult to obtain multiple autonomous system numbers from the Internet. Private autonomous system numbers would have to be used, which would make connections into the Internet complex.

Scenario 16-2

The ISP Interconnect Corp. is a startup company that is configuring its network. The company has a well-resourced network and is in the process of configuring the iBGP within the autonomous system. Figure 16-12 provides the diagram for the network in this scenario.

Figure 16-12 Diagram for Scenario 16-2



1. The original design required a fully meshed iBGP network. This was calculated to mean 250 connections, which was deemed unacceptable because it would be too great of a drain on resources. Route reflectors are obviously the answer. Configure Router A to run BGP and act as a route reflector to clients B and C.
2. Configure Routers B and C to run BGP as clients to Router A.
3. Having configured the cluster, are any other tasks necessary?
4. The company has decided in its early stages to require organizations connecting into it to use default routes. How would these routes be disseminated without the organization's autonomous system, and how would an interior router running only an IGP determine which path out of the routing domain to take if it had more than one connection?
5. Given that the use of a default route by the client organization gives it the least configuration power to manage and manipulate its traffic flow, how would the network administrator justify this simple approach?

Scenario 16-3

Review output screens in Examples 16-10 and 16-11, and answer the questions that follow.

Example 16-10 Scenario 16-3 Output Screen 1

```
Router# show ip prefix-list detail tryout
ip prefix-list tryout:
Description: tryout-list
count: 1, range entries: 0, sequences: 5 - 10, refcount: 3
seq 5 permit 130.0.0.0/8 (hit count: 0, refcount: 1)
seq 10 permit 44.0.0.0/8 (hit count: 28, refcount: 1)
```

Example 16-11 Scenario 16-3 Output Screen 2

```
Router# show ip bgp
BGP table version is 457017, local router ID is 200.172.136.21
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop        Metric   LocPrf Weight Path
*> 7.2.0.0/16       144.39.228.49    0       100   701  I
*> 17.14.96.0/24    144.39.128.7     0       100   99  17I
*> 17.4.126.0/24    144.39.228.49    0       100   701  I
*> 17.96.91.0/24    144.39.228.49    0       100   701  I
*> 33.48.2.0/23     144.39.228.49    0       100   701  I
*> 33.48.7.0/24     144.39.228.49    0       100   701  I
*> 33.48.8.0/24     144.39.228.49    0       100   701  I
*> 33.48.14.0/24    144.39.228.49    0       100   701  I
*> 33.48.123.0/24   144.39.228.49    0       100   701  i
```


Example 16-11 *Scenario 16-3 Output Screen 2 (Continued)*

```

*> 33.48.124.0/23    144.39.228.49      0    200  701 I
*> 33.48.124.0/23    144.39.128.7       0    100  99 17 i
*> 33.48.176.0/20    144.39.228.49      0    100  701 i
*> 33.49.32.0/23     144.39.228.49      0    100  701 i
*> 33.49.35.0/24     144.39.228.49      0    100  701 i
*> 33.49.36.0/23     144.39.228.49      0    100  701 i
*> 33.96.0.0/18      144.39.228.49      0    100  701 i
*> 33.143.24.0/21    144.39.228.49      0    100  701 i
*> 33.154.0.0/18     144.39.228.49      0    100  701 i
*> 33.216.0.0/24     144.39.228.49      0    100  701 i
*> 33.216.0.0/19     144.39.228.49      0    150  701 i

```

1. Using Example 16-10, identify how many times the route 44.0.0.0/8 has been sent in outgoing updates from the router.
2. Which path will be chosen in Example 16-11 to get to 33.48.124.0, and why?
3. What is the most likely reason for the source of a route to be flagged as incomplete?
4. To send packets to network 7.2.0.0/16, the router will direct traffic to a next-hop router. The data frame at Layer 2 will be addressed to this next hop, which will route it on to the next router in the journey to its destination. What is the Layer 3 address of the next logical hop, and why was it selected?

Scenario Answers

The answers provided in this section are not necessarily the only possible answers to the questions. The questions are designed to test your knowledge and to give practical exercise in certain key areas. This section is intended to test and exercise skills and concepts detailed in the body of this chapter.

If your answer is different, ask yourself whether it follows the tenets explained in the answers provided. Your answer is correct not if it matches the solution provided in the book, but rather if it has included the principles of design laid out in the chapter.

In this way, the testing provided in these scenarios is deeper: It examines not only your knowledge, but also your understanding and ability to apply that knowledge to problems.

If you do not get the correct answer, refer back to the text and review the subject tested. Be certain to also review your notes on the question to ensure that you understand the principles of the subject.

Scenario 16-1 Answers

1. *Give reasons to support Humugos's desire to have multiple connections to the Internet.*

Multiple connections to the Internet not only would provide redundancy, but also could be configured to load balance traffic into the Internet. If load balancing is not an option because the multiple connections are to different ISPs, traffic management could still be enforced by using each link for different purposes. Tuning the attributes and configuring prefix lists would do this very effectively.

2. *The company has been advised to redistribute static routes into the Internet BGP. It had intended to redistribute dynamic OSPF routes directly into the ISP provider. Explain why the ISP was not in favor of this configuration.*

If the ISP accepted routes that had been dynamically redistributed into its autonomous system from OSPF, it could have a very unstable network. The problem is that every time there is a change anywhere that results in an update being generated by OSPF, it is redistributed into BGP, requiring BGP to process this change and generate an update. The probability is that no aggregation is configured, which leads to additional traffic and large routing tables. The last problem is that any error experienced by OSPF propagates into BGP and can cause unstable routing tables.

3. *Using Figure 16-11, issue the configuration commands that would allow Router B connecting into the Internet to select the path to network 78.9.80.0 via Router G. Use the local preference attribute to select the path.*

Given the design of the network, the path to network 78.9.80.0 has a longer AS_Path through Router G. To tune the local preference to select this path means altering the selection that it would naturally have taken. The configuration commands are as follows:

```
Router B(config)#router bgp 100
network 167.55.0.0 mask 255.255.0.0
neighbor 131.171.67.8 remote-as 100
neighbor 201.101.9.8 remote-as 250
bgp default local-preference 250
```

Remember that the higher the preference, the more likely the selection.

Scenario 16-2 Answers

1. *The original design required a fully meshed iBGP network. This was calculated to mean 250 connections, which was deemed unacceptable because it would be too great of a drain on resources. Route reflectors are obviously the answer. Configure Router A to run BGP and as a route reflector to clients B and C.*

```
router bgp 100
network 10.0.0.0
network 131.108.0.0
neighbor 10.10.10.10 remote-as 100
neighbor 10.10.10.10 route-reflector-client
neighbor 10.10.8.9 remote-as 100
neighbor 10.10.8.9 remote-as 100 route-reflector-client
```

2. *Configure Routers B and C to run BGP as clients to Router A.*

Router B:

```
router bgp 100
network 10.0.0.0
neighbor 10.10.10.9 route-reflector-client
```

Router C:

```
router bgp 100
network 10.0.0.0
neighbor 10.10.8.10 remote-as 100
```

3. *Having configured the cluster, are any other tasks necessary?*

Given that the route reflector is now forwarding the routes between B and C, the link between these routers is no longer necessary, and the BGP link between them should be broken. This simply requires the removal of the neighbor statements that create the link on both Routers B and C.

4. *The company has decided in its early stages to require organizations connecting into them to use default routes. How would these routes be disseminated without the organization's autonomous system, and how would an interior router running only an IGP determine which path out of the routing domain to take if it had more than one connection?*

The routers in the client organization do not need to run BGP. They simply need to configure a default route and propagate this into the routing domain, in accordance with the interior routing protocol that is being run.

If the autonomous system were multihomed into the Internet, there would be more than one default route propagated throughout the system. Any router within the autonomous system would determine the best path to the outside world by comparing the routing protocol metrics between the default routes. Thus, RIP would select the lowest hop count, EIGRP the lowest combination of bandwidth and delay, and OSPF the lowest cost.

5. *Given that the use of a default route by the client organization gives it the least configuration power to manage and manipulate its traffic flow, how would the network administrator justify this simple approach?*

The default route, although giving the least control over the connection to the Internet, is very robust in that it has no working parts to fail. Therefore, it requires very little CPU or memory. The lack of redistribution eliminates the possibility of routing loops, and the lack of a routing protocol running over the physical link to the Internet frees up bandwidth for data.

Scenario 16-3 Answers

1. *Using Example 16-10, identify how many times the route 44.0.0.0/8 has been sent in outgoing updates from the router?*

The prefix list tryout has 28 hits logged for the network 44.0.0.0/8. This means that 28 updates have been sent with the network 44.0.0.0 from the router to its neighbors.

2. *Which path will be chosen in Example 16-11 to get to 33.48.124.0, and why?*

The path using 144.39.228.49 as the next hop will be used. The local preference is set to 200, because the local preference prefers a higher value.

3. *What is the most likely reason for the source of a route to be flagged as incomplete?*

The route was probably redistributed into BGP, and it therefore cannot identify as much information as if it were received as a routing update with attributes attached.

4. *To send packets to network 7.2.0.0/16, the router will direct traffic to a next-hop router. The data frame at Layer 2 will be addressed to this next hop, which will route it on to the next router in the journey to its destination. What is the Layer 3 address of the next logical hop, and why was it selected?*

The next logical hop for the route 7.2.0.0/16 is 144.39.228.49. This address was selected because it is the next hop in the best path to the destination. BGP determined the best path based on AS_Path. The alternate route has to journey through two autonomous systems to find the destination network, so this path has a more direct route. Because neither the weight attribute nor the local preference attribute has been tuned, the AS_Path is the determining attribute. This information is not shown in the output screen.

Part VI: Redistribution and Policy-Based Routing

Chapter 17 Implementing Redistribution and Controlling Routing Updates

Chapter 18 Controlling Network Traffic with Route Maps and Policy-Based Routing

Part VI covers the following Cisco BSCI exam topics:

- Identify the steps to select and configure the different ways to control routing update traffic
- Identify the steps to configure policy-based routing using route maps
- Identify the steps to configure router redistribution in a network
- Explain the use of redistribution between BGP and Interior Gateway Protocols (IGPs)
- Identify the steps to verify route redistribution
- Interpret the output of various **show** and **debug** commands to determine the cause of route selection errors and configuration problems



This chapter covers the following topics, which you need to understand to pass the CCNP/CCDP/CCIP BSCI exam:

- Understanding the fundamentals of redistribution
- Identifying the issues with redistribution
- Understanding the routing decisions that affect redistribution
- Controlling routing updates during redistribution
- Configuring redistribution
- Configuration commands to control routing updates in redistribution
- Controlling routing updates with filtering
- Verifying, maintaining, and troubleshooting the implementation of redistribution and filtering

Implementing Redistribution and Controlling Routing Updates

The topics in this chapter deal with the traffic generated by routing updates in terms of both the network resources that they use and the information contained within them. This covers two different but related areas, redistribution and filtering. The network overhead involved in routing updates has already been dealt with in other chapters. It keeps recurring as a theme because all network traffic directly influences the network's capability to scale or to grow.

The information propagated through the network is complex when dealing with one routing protocol. When multiple protocols have to share information (through redistribution) so that the larger network can see every route available within the autonomous system, the information flow must be controlled and managed very closely with filtering and other solutions.

This chapter deals with the need for redistribution, which increases the network overhead, and filtering, which is used to reduce overhead. The chapter explains the design issues that might affect the configuration, followed by configuration examples of route redistribution and filtering.

“Do I Know This Already?” Quiz

The purpose of the “Do I Know This Already?” quiz is to help you decide what parts of this chapter to use. If you already intend to read the entire chapter, you do not necessarily need to answer these questions now.

The 14-question quiz, derived from the major sections in the “Foundation Topics” portion of the chapter, helps you to determine how to spend your limited study time.

Table 17-1 outlines the major topics discussed in this chapter and the “Do I Know This Already?” quiz questions that correspond to those topics.

Table 17-1 “Do I Know This Already?” Foundation Topics Section-to-Question Mapping

Foundation Topics Section	Questions Covered in This Section
Understanding the Fundamentals of Redistribution	1–4
Understanding the Routing Decisions That Affect Redistribution	5, 6
Controlling Routing Updates During Redistribution	7
Configuring Redistribution	8–11
Controlling Routing Updates with Filtering	12, 13
Verifying, Maintaining, and Troubleshooting the Implementation of Redistribution and Filtering	14

CAUTION The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark this question wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

1. Which of the following are reasons to run multiple routing protocols?
 - a. Applications requiring UNIX
 - b. Smaller routing domains speed up convergence
 - c. Political enclaves
 - d. Smaller routing domains ensure a more stable network

2. EIGRP automatically redistributes into which routing protocols?
 - a. IGRP running the same AS number
 - b. IGRP with any process ID
 - c. EIGRP running the same AS number
 - d. EIGRP with any process ID

3. What do the letters SIN represent?
 - a. Simple Internetwork
 - b. Ships in the night

- c. Structured IP Network
 - d. Single IP Network
4. How many IP routing tables can be held on a router within a typical organization?
 - a. One
 - b. One per routing protocol
 - c. Four
 - d. Six
 5. The problems experienced as a result of multiple routing processes and their redistribution include which of the following?
 - a. Suboptimal path
 - b. Loss of Hello packets
 - c. Routing loops
 - d. Continuous LSA propagation
 6. What action is taken if no seed or default metric is configured for OSPF when redistributing EIGRP?
 - a. The route is not entered into the routing table.
 - b. The route is entered with a cost of 0.
 - c. The route is read into OSPF with a cost of 20 (type E2).
 - d. The route is entered with a cost of 20 (type 1).
 7. What techniques can be employed to avoid redistribution problems?
 - a. Distribute lists
 - b. Change administrative distance
 - c. Ensuring the default metric is set to 0
 - d. Redistributing on all border routers in both directions
 8. What is the purpose of distribute lists?
 - a. Determine the administrative distance of a distributed routing protocol
 - b. Identify which interfaces will send updates
 - c. Determine which networks are sent in updates
 - d. Determine which networks are accepted into the routing table

9. Where are distribute lists defined?
 - a. At the interface
 - b. Under the routing process
 - c. At the router level
 - d. At the executive prompt
10. Which command is used to establish the default or seed metric for EIGRP?
 - a. **default-metric 5**
 - b. **metric** *bandwidth delay reliability loading mtu*
 - c. **default-metric** *bandwidth delay reliability loading mtu* **eigrp**
 - d. **default-metric** *bandwidth delay reliability loading mtu*
11. Which command is used to configure the administrative distance?
 - a. **administrative distance**
 - b. **distance**
 - c. **ip default-distance**
 - d. **ip administrative distance**
12. Why might it be necessary to control the routing updates?
 - a. Security
 - b. Prevention of routing loops
 - c. Scaling the network
 - d. Preserving the metric
13. All of the following statements are true; however, which of the following actions is taken first?
 - a. Do not advertise the route if it is matched by a **deny** statement.
 - b. If no match is found in the distribute list, the implicit **deny any** at the end of the access list will cause the update to be dropped.
 - c. If a filter is present, the router examines the access list to see if there is a match on any of the networks in the routing update.
 - d. Advertise the route if matched by a **permit** statement.

14. Which commands could be used to verify and troubleshoot a network that is redistributed?
- a. **show ip protocol**
 - b. **show ip route**
 - c. **show ip route** *routing-protocol*
 - d. **show redistributed**

The answers to this quiz are found in Appendix A, “Answers to Chapter ‘Do I Know This Already?’ Quizzes and Q&A Sections.” The suggested choices for your next step are as follows:

- **7 or less overall score** —Read the entire chapter. This includes the “Foundation Topics” and “Foundation Summary” sections, the “Q&A” section, and the “Scenarios” at the end of the chapter.
- **8–11 overall score** —Begin with the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. If you have trouble with these exercises, read the appropriate sections in “Foundation Topics.”
- **12 or more overall score** —If you want more review on these topics, skip to the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. Otherwise, move to the next chapter.

Foundation Topics

Understanding the Fundamentals of Redistribution

It is rare to find just one routing protocol running within an organization. If the organization is running multiple routing protocols, you need to find some way of passing the networks learned by one routing protocol into another so that every workstation can reach every other workstation. This process is called *redistribution*.

Redistribution is used when a router is receiving information about remote networks from various sources. Although all the networks are entered into the routing table and routing decisions are made on all the networks present in the table, a routing protocol propagates only those networks that it has learned through its own process. When there is no sharing of network information between the routing processes, it is referred to as *ships in the night (SIN)* routing.

Redistribution is often necessary within a network, if only as a transitional implementation. Nonetheless, it should not be thought of as a quick and easy solution. Although route redistribution is often a lifesaver for your network, it is fraught with complexity. Understanding the operation of the processes that you have implemented and how this influences your network is crucial. This chapter focuses on the main topics dealing with the implementation of redistribution.

Although an organization might have many routing protocols running within its autonomous system, each interior routing protocol sees itself as the only interior routing protocol within the autonomous system. When an interior routing protocol such as EIGRP has routes redistributed into its routing process, it assumes that these routes are from another autonomous system and are therefore external routes. This affects the route selection made by the routing process, and EIGRP prefers the interior routes.

The exterior routing protocols see the organization as the autonomous system that connects to the Internet or a service provider.

In Figure 17-1, the routing table for Router B has entries from RIP and OSPF. There are no entries for EIGRP because this is a single network directly connected to the router. You can see that the RIP updates sent out the interfaces do not include networks from OSPF. There are no entries for EIGRP.

Furthermore, Router C has only connected routes in the routing table. This is because, although EIGRP has been configured, Router C is a stub router. When the other interfaces are configured with addresses and the rest of the EIGRP network is connected to Router C, the network will be populated

Some routing protocols automatically exchange networks, although others require some level of configuration. Table 17-2 shows the subtleties of automatic redistribution.

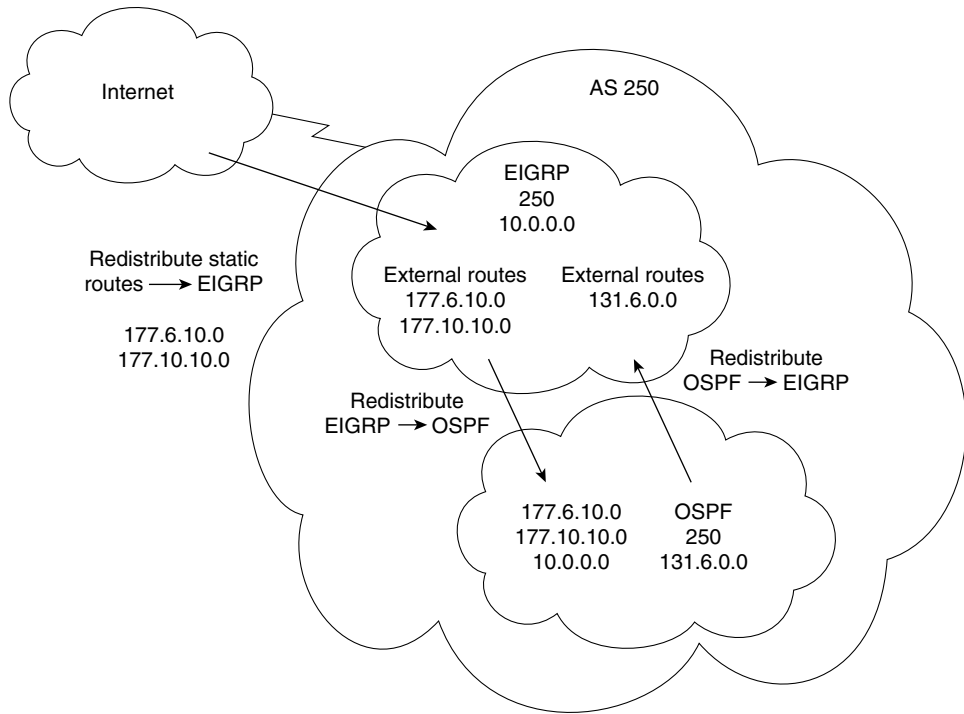
Table 17-2 *Automatic Redistribution Between Routing Protocols*

Routing Protocol	Redistribution Policy
Static	Requires manual redistribution into other routing protocols.
Connected	Unless included in the network command for the routing process, requires manual redistribution.
RIP	Requires manual redistribution.
IGRP	Will automatically redistribute between IGRP and EIGRP if the autonomous system number is the same. Otherwise, it requires manual redistribution.
EIGRP	Will automatically redistribute between IGRP and EIGRP if the autonomous system number is the same. Otherwise, it requires manual redistribution. EIGRP for AppleTalk will automatically redistribute between EIGRP and RTMP. EIGRP for IPX will automatically redistribute between EIGRP and IPX RIP/ SAP; in later versions, NLSP can be manually redistributed.
OSPF	Requires manual redistribution between different OSPF process IDs and routing protocols.
IS-IS	Requires manual redistribution between different routing protocols
BGP	Requires manual redistribution between different routing protocols

Figure 17-2 illustrates redistribution within an organization.

The main reasons for multiple protocols existing within an organization are as follows:

- The organization is transitioning from one routing protocol to another because there is a need for a more sophisticated protocol.
- Historically, the organization was a series of small network domains. The company has plans to transition to a single routing protocol in the future.
- Some departments might have host-based solutions that require different protocols. For example, some UNIX hosts use RIP to discover gateways.
- Often after a merger or a takeover, it takes planning, strategy, and careful analysis to determine the best overall design for the network.

Figure 17-2 *Autonomous Systems Within an Organization*

- Politically, there are ideological differences among the different network administrators, which until now have not been resolved.
- In a very large environment, the various domains might have different requirements, making a single solution inefficient. A clear example is in the case of a large multinational corporation, where EIGRP is the protocol used at the access and distribution layers, but BGP is the protocol connecting the core.

Understanding the Routing Decisions That Affect Redistribution

When embarking on running multiple routing protocols within your network and making one cohesive whole network, redistribution is the answer, but only after careful consideration has been given to the problems that might arise. In order to do this, you need to consider briefly the routing protocol operation, in particular how a path is selected to go into the routing table. For a detailed discussion on routing tables, refer to Chapter 1, “IP Routing Principles.” Path selection is dealt with in depth in Chapter 4, “IP Distance Vector Routing Principles.”

Routing Metrics and Redistribution

There are many routing protocols for IP, and each routing protocol uses a different metric. If the different protocols want to share information through redistribution, the configuration must translate the metrics. The configuration commands are dealt with in the section “Configuring Redistribution,” later in this chapter.

Problems arise when the metrics are redistributed without additional configuration. The metric has no point of reference in the new routing protocol; for example, RIP would be baffled by the metric presented as 786, when expecting a hop count between 0–15. In accepting the new networks, the receiving process must have a starting point, or *seed metric*, in order to calculate the metric for the routing protocol.

The seed metric is assigned to all the routes received into a process through redistribution. The metric is incremented from that point on as the networks propagate throughout the new routing domain.

There are defaults for the seed metrics, but depending on the routing protocol, the default might prevent the route from entering the routing table. The seed metrics are as defined in Table 17-3.

Table 17-3 *Default Seed Metrics*

IP Routing Protocol	Default Seed Metric	Action
RIP	From Cisco IOS release 12.1, the seed metric is infinity	No routes entered into the routing table
IGRP	Infinity	No routes entered into the routing table
EIGRP	Infinity	No routes entered into the routing table
IS-IS	0	Routes entered into the routing table
OSPF	20 (type 2), but routes from BGP are given a metric of 1 (type 2)	Routes entered into the routing table
BGP	MED is given the IGP metric value	Routes entered into the routing table

Remember, the metric is the main method of route selection within a routing protocol. Therefore, it is necessary to define a default seed metric for the networks accepted from the other routing protocol.

Path Selection Between Routing Protocols

Now that route selection within a routing protocol has been explained, this section discusses path selection between routing protocols when more than one routing protocol is running on your network. If the protocols have paths to the same remote destination network, the routing process must decide which path to enter into the routing table. Because the metrics differ between the protocols, selection based on the metric is ruled out as a solution. Instead, another method was devised to solve the problem, the *administrative distance*, as discussed in Chapter 4.

The distinction between the two selection processes is simple: Administrative distance determines between IP routing protocols, and the metric chooses between paths from one routing protocol.

Administrative distance and metrics appear to solve all your problems, until you start to redistribute the information between routing protocols, and the routing process becomes confused as to from where the information came. When the carefully determined rules of selection become tangled, suboptimal routing decisions and routing loops result.

It is therefore important to consider the following rules when redistributing between IP routing protocols:

- If more than one routing protocol is running on a router, the routing table will place the route with the best administrative distance into the routing table.
- In order to be redistributed, the route must exist in the routing table under the ownership of the routing protocol that is being redistributed. Thus, if RIP is being redistributed into EIGRP, the routing table must have an entry for the RIP network.
- When a route is redistributed, it inherits the default administrative distance of the new routing protocol.
- When a route is redistributed, it is considered as an external route to the new routing protocol. For EIGRP and BGP, this means it will inherit the administrative distance of an external route to the new routing protocol. OSPF tracks the route as external and chooses internal routes first.

It is clear that redistribution is not the optimum network design. The simpler and more straightforward the design, the better managed and more stable the network, with fewer errors and faster convergence. Therefore, a hierarchical IP addressing scheme designed to allow continued network growth, combined with a single IP routing protocol that has the scope to support growth, results in a strong, reliable, and fast network. However, it is rare to find a network of any size that runs only one IP routing protocol. When multiple protocols are running, it is necessary to redistribute.

Although the concept of redistribution is straightforward, the design and implementation are extremely tricky. Without a full documented understanding of both the network and the traffic flow, the implementation of redistribution can result in routing loops or the selection of suboptimal paths.

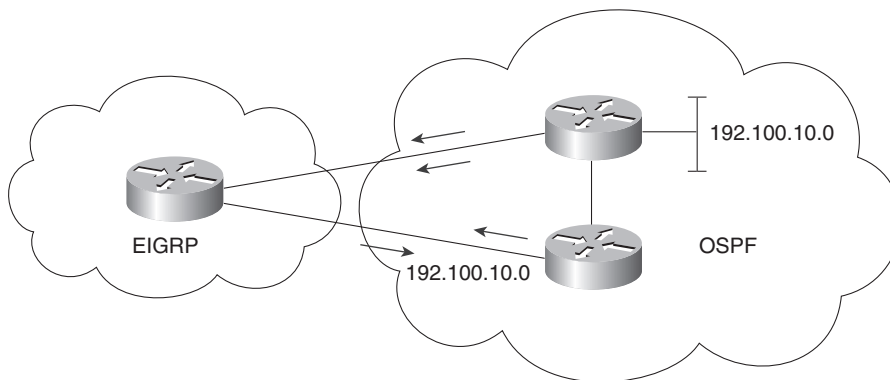
The problems that can occur from redistribution are typically difficult to troubleshoot because the symptoms often appear some distance from the configuration error. The problems experienced as a result of multiple routing processes and their redistribution include the following:

- The wrong, or less efficient, routing decision is made because of the difference in routing metrics. The choice of the less efficient route is referred to as choosing the *suboptimal path*.
- A routing loop occurs, in which the data traffic travels in a circle without ever arriving at the destination. This is normally due to routing feedback, where routers send routing information received from one autonomous system back into the same autonomous system.
- The convergence time of the network increases because of the different technologies involved. If the routing protocols converge at different rates, this might result in timeouts and the temporary loss of networks.
- The decision-making process and the information sent within the protocols might be incompatible and not easily exchanged, leading to errors and complex configuration.

Avoiding Routing Loops When Redistributing

Routing loops occur when a routing protocol is fed its own networks. The routing protocol might see a network as having a more favorable path, although this path points in the opposite direction, into a different routing protocol domain. The potential for confusion is enormous, and it is very easy to create routing loops when redistributing, as shown in Figure 17-3.

Figure 17-3 *How Route Feedback Can Cause Routing Loops*



This problem is solved by the following configurations:

- Changing the metric
- Changing the administrative distance

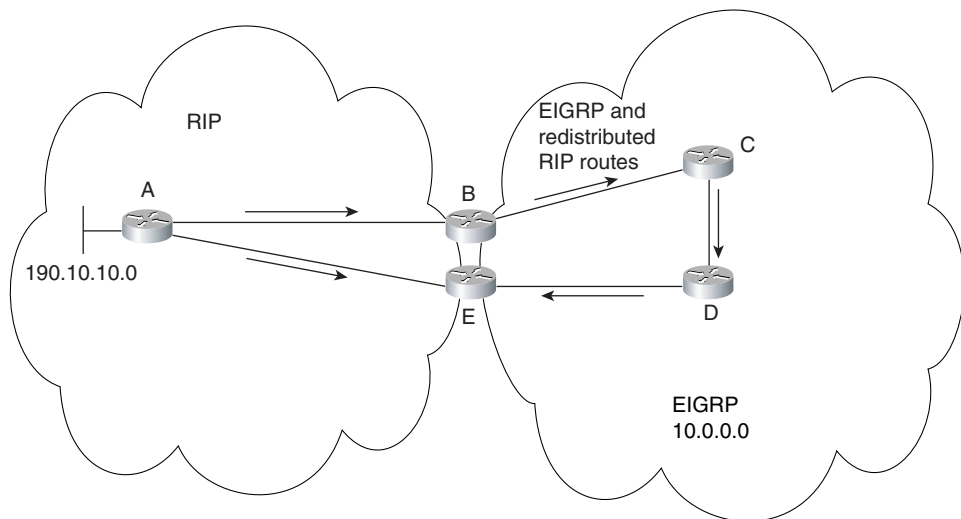
- Using default routes
- Using passive interfaces with static routes
- Using distribute lists

These configurations are discussed in the section of the chapter titled “Configuring Redistribution.”

To manage the complexity of these networks and to reduce the possibility of routing loops, some level of restriction in the information sent across the various domains is often necessary. This is done via filtering, using access lists.

Consider the problem by looking at the example in Figure 17-4, remembering that administrative distance is considered without any reference to the metrics. Imagine for a moment that Router A is running RIP and advertising network 190.10.10.0 to both Routers B and E. When Router B receives the RIP update, it redistributes the network 190.10.10.0 into OSPF and advertises it to Router C, which advertises the network to Router D. Eventually Router E receives an OSPF update from D, reporting a network 190.10.10.0 with the path D, C, B, A. However, Router E has a direct path to Router A via RIP, which would be the preferable path. In this instance, the administrative distance works against the network. Because OSPF has an administrative distance of 110 and RIP has an administrative distance of 120, the path placed in the routing table is the one advertised by OSPF via D, C, B and A. In this case, manually configuring the administrative distance on Routers B and E is advisable.

Figure 17-4 Path Selection Using Administrative Distance



If EIGRP is running on Routers B, C, D, and E, there should be no problems. When RIP redistributes into EIGRP on Router B and the update is propagated to Router E, the routing table should select the route to 190.10.10.0 via Router A. The reason is that when network 190.10.10.0 is redistributed into EIGRP, it is flagged as an external route. Thus, it has the administrative distance of 170 and is discarded in favor of RIP administrative distance of 120. The routing table contains RIP's path to the network 190.10.10.0.

When EIGRP then redistributes into RIP, the routing table, having no EIGRP entry for the network 190.10.10.0, cannot redistribute this network back into the RIP process. Theoretically, a routing loop is avoided. However, in practice this might not be the case, as it is dependent on when the routing updates come into the routing process and the inherent stability of the network. You should avoid two-way redistribution between routing protocols for these reasons, unless you take great care in the design of the network and place filters on the redistributing routers in order to prevent routing protocol feedback.

Remember that although you can change the defaults for administrative distance, you should take care when subverting the natural path selection, and any manual configuration must be done with careful reference to the network design of the organization and its traffic flow. To change the administrative distance of a routing protocol, use the following command:

```
Router(config-router)#distance weight [network wildcard-mask]
```

For a static route use the following command:

```
Router(config)#ip route network [mask] {address | interface} [distance]
```

There are additional commands to change specific types of routes on a per-protocol basis. For example, it is possible to change the administrative distance of internal or external EIGRP routes. This can also be done for OSPF and BGP. To find more on this subject, search for the keyword “distance” on the Cisco web site.

The administrative distance reflects the preferred choice. The defaults are listed in Chapter 4 in Table 4-2.

Avoiding Suboptimal Routing Decisions When Redistributing

Routing loops are only one problem that can result from redistributing routes between routing protocols. As mentioned in the previous section, suboptimal routing is sometimes created by redistribution. For example, the administrative distance selects the suboptimal path when a directly connected network is designed as a backup link. Although this is a problem of administrative distance as opposed to redistribution, it is important to ensure that the suboptimal path is not propagated into the new routing protocol. To overcome this problem—the administrator's preferred route not coinciding with that of the routing protocol—a floating static route is configured, as described in Chapter 4.

The following are guidelines to keep in mind when designing your network to avoid routing loops and suboptimal path selection when redistributing between routing protocols:

- Have a sound knowledge and clear documentation of the following:
 - The network topology (physical and logical)
 - The routing protocol domains
 - The traffic flow
- Do not overlap routing protocols. It is much easier if the different protocols can be clearly delineated into separate domains, with routers acting in a similar function to area border routers (ABRs) in OSPF. This is often referred to as the *core and edge protocols*.
- Identify the boundary routers on which redistribution is to be configured.
- Determine which protocol is the core and which protocol is the edge protocol.
- Determine the direction of redistribution, that is, into which routing protocol the routes are to be distributed.
- If redistribution is needed, ensure that it is a one-way distribution, where the one routing protocol redistributes into another routing protocol, but the other routing protocol does not redistribute back. For example, RIP redistributes into EIGRP, but EIGRP does not redistribute into RIP. This avoids networks being fed back into the originating domain. Use default routes to facilitate the use of one-way redistribution, if necessary.
- If two-way redistribution cannot be avoided, use the mechanisms in the following list:
 - Manually configuring the metric
 - Manually configuring the administrative distance
 - Using distribution access lists

Avoiding Problems with Network Convergence When Redistributing

To maintain consistent and coherent routing among different routing protocols, you must consider the different technologies involved. A major concern is the computation of the routing table and how long it takes the network to converge. Although EIGRP is renowned for its speed in convergence, RIP has a poorer reputation in this regard. Sharing the network information across the two technologies might cause some problems.

For example, the network converges at the speed of the slower protocol. At some point, this will create timeouts and possibly routing loops. Adjusting the timers might solve the problems, but any routing protocol configuration must be done with a sound knowledge of the entire network and of the routers that need to be configured. Timers typically require every router in the network to be configured to the same value.

Controlling Routing Updates During Redistribution

Controlling routing updates is useful for many reasons. The reasons for controlling routing updates include:

- To hide certain networks from the rest of the organization
- To prevent routing loops
- To control the network overhead or traffic on the wire, allowing the network to scale
- For simple security reasons

Various methods enable you to control the routing information sent between routers during redistribution. These methods include the following:

- Passive interfaces
- Static routes
- Default routes
- The null interface
- Distribute lists
- Route maps

The next sections describe each method in more detail.

Passive Interfaces

A passive interface does not participate in the routing process. In RIP and IGRP, the process listens but will not send updates. In OSPF and EIGRP, the process neither listens nor sends updates because they do not send Hellos, and therefore no neighbor relationship can form.

The interfaces that participate in the interior routing process are controlled by the interface configuration. During configuration, the routing process is instructed via the **network** command on which interfaces to use. Because most protocols express the networks at the major boundary, interfaces that have no reason to send this protocol's updates propagate the data across the network. This is not only a waste of bandwidth, but in many cases, it can lead to confusion, particularly during redistribution. The configuration of passive interfaces to prevent updates going into the domains of other routing protocols can simplify the network administration and prevent routing loops.

Static Routes

A static route is a route that is manually configured. It takes precedence over routes learned by a routing process because it has a lower default administrative distance.

If no routing process is configured, static routes can be configured to populate the routing table. This is not practical in a large network because the table cannot learn of changes in the network topology dynamically. In small environments or for stub networks, however, this is an excellent solution. It is used to good effect when there are multiple protocols configured. Instead of redistributing the entire routing tables between the protocols, static routes are defined and redistributed. This is useful if you need to provide more information than a default route. The routing protocols have the information they need while you maintain careful control over the design and data flow. Again, this is a typical scenario for BGP and an IGP to exchange information.

The reasons for static routing are summarized as follows:

- To prevent the need for a routing protocol to run on the network, reducing the network overhead to zero. This can be used with dialup lines (dial-on-demand routing).
- If there are two autonomous systems that do not need to exchange the entire routing table, but simply need to know about a few routes.
- No routing protocol is configured, for example, on a remote stub node.
- To change the mask of the network. For example, as seen in BGP, you can statically define a supernet and redistribute the static route into the BGP process. This is also done when redistributing from a routing protocol that understands VLSM to one that does not.

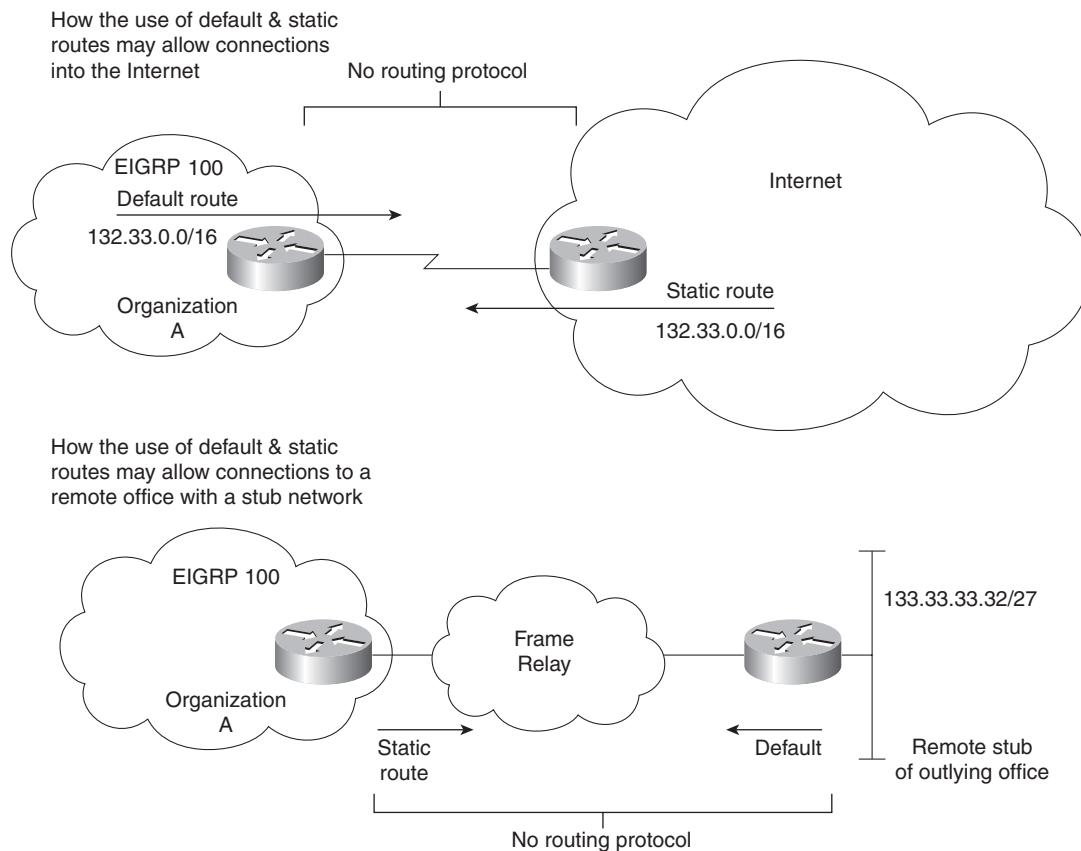
Default Routes

A default route is used if there is no entry in the routing table for the destination network. If the lookup finds no entry for the desired network and no default route is configured, the packet is dropped.

If the routing process is denied the right to send updates, the downstream routers will have a limited understanding of the network. To resolve this, use default routes. Default routes reduce overhead, add simplicity, and can remove loops, particularly when used instead of redistribution between routing protocols. One routing protocol can use a default route to the other routing protocol's domain; a typical example would be an IGP pointing a default route into the BGP domain.

Another occasion for configuring a default route would be for a stub network to connect to the larger network.

Default and static routes are shown in Figure 17-5.

Figure 17-5 *The Use of Default and Static Routes*

The null Interface

The null interface is an imaginary interface that is defined as the next logical hop in a static route. All traffic destined for the remote network is carefully routed into a black hole. This is used to good effect in redistribution, as it is used either to discard routes by destination in a rudimentary filtering system or to redistribute between classless and classful routing protocols.

It is used to feed routes into the other routing protocol, allowing another mask to be set. In this way, it aggregates routes as shown in Chapter 16, “Implementing and Tuning BGP for Use in Large Networks.”

Distribute Lists

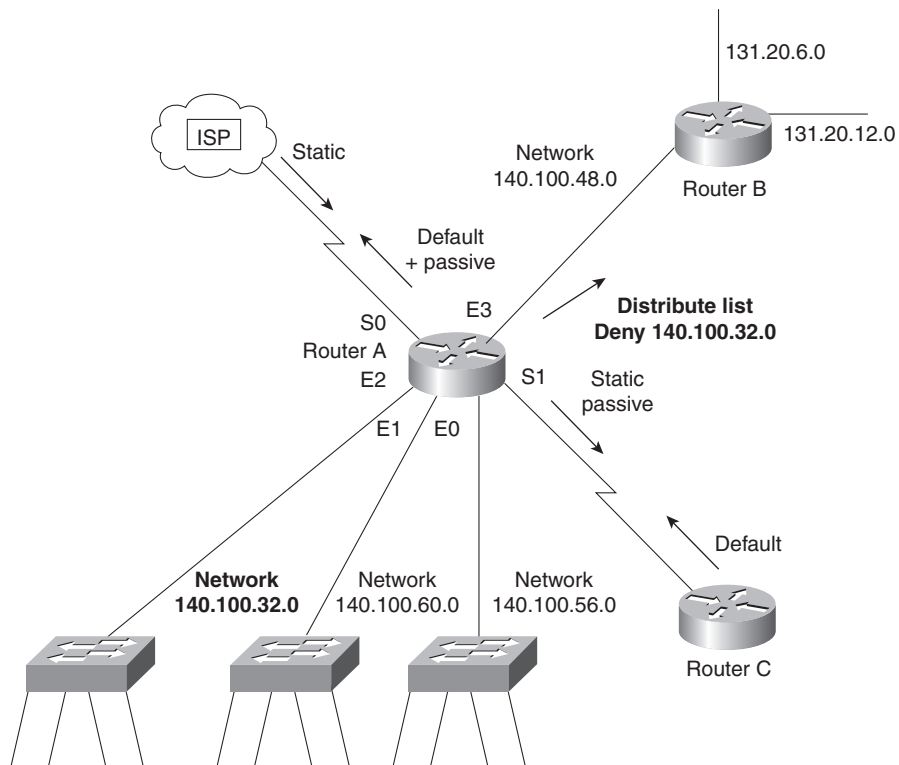
Distribute lists are access lists applied to the routing process, determining which networks are accepted into the routing table or sent in updates. When communicating to another routing process through redistribution, it is important to control the information sent into the other process. This control is for security, overhead, the prevention of routing loops, and management reasons. Access lists afford the greatest control for determining the traffic flow in the network.

Route Maps

Route maps are complex access lists that permit conditional programming. If a packet or route matches the criteria defined in a **match** statement, changes defined in the **set** command are performed on the packet or route in question. These are used in redistribution in the same way as distribute lists, but allow a greater level of sophistication in the criteria stated.

Figure 17-6 shows the options for controlling routing updates in a large and complex network.

Figure 17-6 *Controlling Routing Updates*



In Figure 17-6, Router A has a distribute list that is denying the propagation of the network 140.100.32.0 out of E3, which is the network connected to E2. Network 140.100.32.0 might have some security reasons for not being seen by the networks connected to Router B. This network could be a test or an R&D project for the department connecting to Router B, and connectivity would confuse the situation.

S0 and S1 have static routes configured. In the case of S0, this is the connection into the Internet, and the static routes are configured by the ISP. This allows them to connect to the ISP without having to receive dynamic routing updates from the ISP. The routing updates from the ISP containing the Internet routing tables could be huge.

The organization has a default route set so that anyone wanting to flee the organization's network can send to the default route, thus keeping the routing tables small and the update traffic to a minimum.

On S1, the router's interface is configured with static routes so that the router at the other end does not need to run a routing protocol. The router at the other end has a default route configured, and the suggestion is that this is a stub network. This ensures that Router C has a simple configuration with few demands on the router.

The concepts of redistribution are detailed in the following examples with configuration scripts. This will reinforce the concepts and understanding of the technology.

Configuring Redistribution

Redistribution configuration is specific to the routing protocol itself. Before you contemplate implementation, reference the configuration guides from Cisco.

All protocols require the following steps for redistribution:

- Step 1** Configure redistribution.
- Step 2** Define the default metric to be assigned to any networks that are distributed into the routing process.

The commands for redistribution are configured as subcommands to the routing process. The **redistribute** command identifies the routing protocol *from* which the updates are to be accepted. It identifies the source of the updates.

These commands, discussed in detail in the next sections, constitute the basic steps in the implementation of redistribution. Depending on the design of your network, additional configuration might be needed. The configuration of administrative distance, passive interfaces, and

static and default routes are provided in the section “Configuration Commands to Control Routing Updates in Redistribution.”

Redistribution Configuration Syntax

To configure redistribution between routing protocols, the following command syntax is used under the routing protocol that receives the routes:

```
Router(config-router)#redistribute protocol [process-id] {level-1 | level-1-2 | level-2}
[metric metric-value] [metric-type type-value]
[match {internal | external 1 | external 2}] [tag tag-value]
[route-map map-tag] [weight weight] [subnets]
```

The command is very complex because it shows all the parameters for all the different protocols. For an explanation of the command parameters, refer to Table 17-4.

Table 17-4 Command Description of Redistribution

Command	Description
<i>protocol</i>	This is the routing protocol that provides the routes. Remember, most commands with two parameters are structured from a value to a value or from a source to a destination. Routes are redistributed from this source protocol. It can be one of the following keywords: connected , bgp , eigrp , egp , igrp , isis , iso-igrp , mobile , ospf , static , or rip .
<i>process-id</i>	For BGP, EGP, EIGRP, or IGRP, this is an autonomous system number. For OSPF, this is an OSPF process ID. RIPv1 and RIPv2 do not use either a process ID or an autonomous system number.
level-1	For IS-IS, Level 1 routes are redistributed into other IP routing protocols independently.
level-1-2	For IS-IS, both Level 1 and Level 2 routes are redistributed into other IP routing protocols.
level-2	For IS-IS, Level 2 routes are redistributed into other IP routing protocols independently.
metric <i>metric-value</i>	This optional parameter is used to specify the metric used for the redistributed route. When redistributing into protocols other than OSPF, if this value is not specified and no value is specified using the default-metric router configuration command, routes cannot be redistributed. If the routes are redistributed, the routing table might become completely confused with mismatched metrics. Although it might be possible to configure networks to be redistributed into another protocol without specifying a metric, it is ill advised. Configure a specific or default metric, using a value consistent with the destination protocol. Remember that you are influencing the path selection made by the routing process.

(continues)

Table 17-4 *Command Description of Redistribution (Continued)*

Command	Description
metric-type <i>type-value</i>	This is an optional OSPF parameter that specifies the external link type associated with the default route advertised into the OSPF routing domain. This value can be 1 for type 1 external routes, or 2 for type 2 external routes. The default is 2. Refer to Chapter 8, “Using OSPF Across Multiple Areas,” for more detail on OSPF external route types.
match	This is an optional OSPF parameter that specifies the criteria by which OSPF routes are redistributed into other routing domains. It can be one of the following: internal—Redistribute routes that are internal to a specific autonomous system. external 1—Redistribute routes that are external to the autonomous system but that are imported into OSPF as a type 1 external route. external 2—Redistribute routes that are external to the autonomous system but that are imported into OSPF as a type 2 external route.
tag <i>tag-value</i>	(Optional) The tag-value is a 32-bit decimal value attached to each external route. It is not used by the OSPF protocol itself, but it can be used to communicate information between autonomous system boundary routers. If no value is specified, then the remote autonomous system number is used for routes from BGP and EGP; for other protocols, zero (0) is used.
route-map	(Optional) This instructs the redistribution process that a route map must be referenced to filter the routes imported from the source routing protocol to the current routing protocol. If it is not specified, all routes are redistributed because no filtering is performed. If this keyword is specified but no route map tags are listed, no routes will be imported. It is important, therefore, to pay attention to the configuration.
<i>map-tag</i>	This is the optional identifier of a configured route map to filter the routes imported from the source routing protocol to the current routing protocol. Route maps are covered later in this chapter.
weight <i>weight</i>	(Optional) This sets the attribute of network weight when redistributing into BGP. The weight determines the preferred path out of a router when there are multiple paths to a remote network. This is an integer between 0 and 65,535.
subnets	(Optional) For redistributing routes into OSPF, this is the scope of redistribution for the specified protocol. It is important to remember that this is required to bring subnets of classful networks.

Configuring the Default Metric

The default metric can be configured in several ways. The first is to include the metric in the **redistribute** command, defining the metric for that specific redistribution, which you saw in the preceding command syntax. You can also configure the default metric with a command defined under the routing process. Using the **default-metric** command saves work because it eliminates the need to define metrics separately for each redistribution.

Example 17-1 shows the metric included in the **redistribute** command.

Example 17-1 Including the Metric in the **redistribute** Command

```
Router(config)#router eigrp 100  
Router(config-router)#redistribute rip metric 10000 100 255 1 1500  
Router(config-router)#network 140.100.0.0
```

This configuration shows the following:

- The use of the **redistribute** command
- The routing process from which the routes are being accepted
- The metric parameter, allowing the configuration of the EIGRP to state the new metric that the old RIP networks will use while traversing the EIGRP network

Configuring the Default Metric for OSPF, IS-IS, RIP, EGP, or BGP

It is possible to redistribute the routing protocol and then, with a separate command, to state the default metric. The advantage is that it is a simpler configuration visually, which is helpful in troubleshooting. Also, if more than one protocol is being redistributed into the routing protocol, the default metric applies to all the protocols being redistributed. IS-IS cannot define a default metric. The metric must be stated when redistributing. If no metric is stated, the default (0 cost) is entered and the route discarded.

To configure the default metric for OSPF, RIP, EGP, or BGP, use the following command syntax:

```
Router(config-router)#default-metric number
```

The **default-metric** command is used in Example 17-2.

Example 17-2 Configuring the Default Metric for Static and OSPF Routes Received by RIP

```
Router(config)#router rip  
Router(config-router)#redistribute static  
Router(config-router)#redistribute ospf 25  
Router(config-router)#default-metric 2
```

In this example, the default metric is set to 2. You are advised to set the metric to a low number so that when RIP increments the metric, it can do so without hitting the upper limit of the metric, 15.

Configuring the Default Metric for EIGRP or IGRP

To configure the default metric for IGRP or EIGRP, use the following command syntax:

```
Router(config-router)#default-metric bandwidth delay reliability loading mtu
```

Typically, you should take the values shown on one of the outgoing interfaces of the router being configured by issuing the following exec command:

```
Router#show interface
```

The significance of the metric values is shown in Table 17-5.

Table 17-5 The Parameters of the default-metric Command

Command Parameter	Description
<i>bandwidth</i>	The minimum bandwidth seen on the route to the destination. It is presented in kilobits/per second (kbps).
<i>delay</i>	The delay experienced on the route and presented in tens of microseconds.
<i>reliability</i>	The probability of a successful transmission given the history of this interface. The value is not used. It is expressed in a number from 0 to 255, where 255 indicates that the route is stable and available.
<i>loading</i>	A number range of 0 to 255, where 255 indicates that the line is 100 percent loaded. This parameter is not used either and is, therefore, set to one.
<i>mtu</i>	The maximum packet size that can travel through the network. This parameter is not used and is therefore set to 1500.

Example 17-3 shows the configuration of the default metric when redistributing between routing protocols.

Example 17-3 Configuring the Default Metric for EIGRP

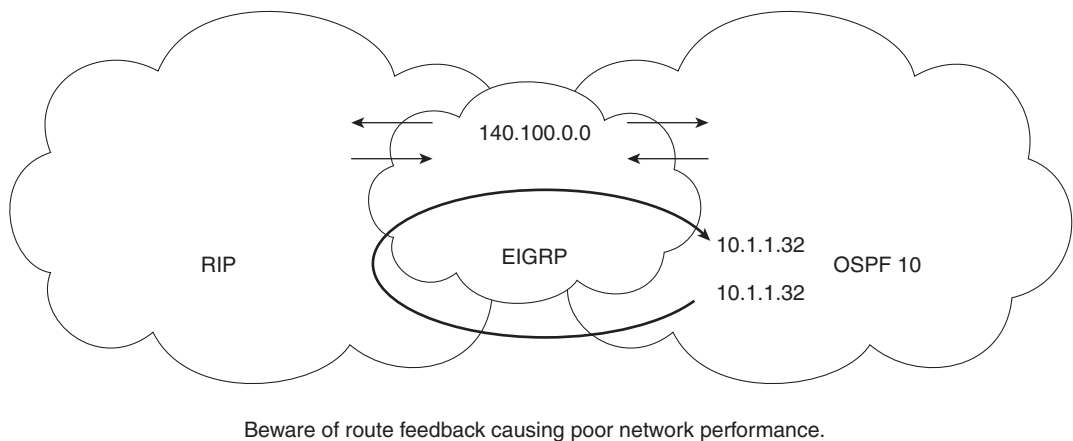
```
Router(config)#router eigrp 100
Router(config-router)#redistribute rip
Router(config-router)#redistribute ospf 10
Router(config-router)#default-metric 10000 100 255 1 1500
Router(config-router)#network 140.100.0.0
```

In Example 17-3, EIGRP assigns networks from both RIP and OSPF the same seed metric. Although this design and configuration seems complex, it is fairly common. Imagine the situation in which OSPF and RIP have been running. The decision to transition the network to EIGRP has been made.

The network designed for EIGRP will run in the core, with the edge routers running redistribution. RIP has been included in the design map to accommodate the UNIX systems running Routed, which is the routing process for UNIX systems.

The default, or seed, metric used is the bandwidth, delay, reliability, load, and maximum transmission unit (MTU), which reflect the compound metric used by IGRP and EIGRP. However, RIP and OSPF would supply in the configuration a number for hop count and cost, respectively. (Refer to Example 17-2.) This design requires careful consideration and filtering of routing updates because it can result in route feedback, as shown in Figure 17-7.

Figure 17-7 *Configuring the Default Metric*



In this figure, EIGRP redistributes the network 10.1.1.32 from OSPF throughout its domain. At the other end of the domain, EIGRP is redistributed into RIP. The network is now known by all routers, whatever their routing protocol preference. However, if another router on the border between the EIGRP domain and the RIP domain is redistributing, it will hear the RIP update for 10.1.1.32 and redistribute all its networks into EIGRP, including the network originating from OSPF. EIGRP will in turn redistribute 10.1.1.32 back into OSPF, creating a classic routing loop.

Configuring the Administrative Distance

As shown in this chapter, it is important to ensure that routes redistributed into another routing protocol are assigned an appropriate metric. However, it is equally important to consider the need to control the choice that the routing process makes when presented with multiple routes to the same destination from different routing protocols. The metric is not appropriate because the multiple routes are from different routing protocols that are not redistributing. Changing the administrative distance allows the best path to be chosen.

To ensure that the optimal path is chosen, it is sometimes necessary to change the administrative distance to make it less favorable. The command structure is protocol-dependent, in that EIGRP requires a separate command. The following command syntax is used for EIGRP:

```
Router(config)#distance eigrp internal-distance external-distance
```

The **distance** command, as used to configure the EIGRP administrative distance, is explained in Table 17-6.

Table 17-6 Configuring Administrative Distance for EIGRP

Command	Description
<i>internal-distance</i>	Administrative distance for EIGRP internal routes. These are routes learned from another entity within the same autonomous system.
<i>external-distance</i>	Administrative distance for EIGRP external routes. These are routes for which the best path is learned from a neighbor external to the autonomous system, such as EIGRP from another autonomous system or another TCP/IP routing protocol, such as OSPF.

To configure the administrative distance for the other IP protocols, the following command syntax is used:

```
Router(config-router)#distance weight [address mask] [access-list-number | name] [ip]
```

The **distance** command to configure the administrative distance for other IP protocols is explained in Table 17-7.

Table 17-7 Configuring Administrative Distance for Other IP Protocols

Command	Description
<i>weight</i>	Administrative distance, an integer from 10 to 255, where 255 means that the route is unreachable. The values 0 to 9 are reserved for internal use.
<i>address</i>	Optional IP address. Assigns the administrative distance to networks according to the IP address of the router supplying the routing information.
<i>mask</i>	Optional wildcard mask for IP address. A bit set to 1 in the mask argument instructs the software to ignore the corresponding bit in the address value.
<i>access-list-number name</i>	Optional number or name of standard access list to be applied to the incoming routing updates. Assigns administrative distance to matching or permitted networks.
ip	Optional. Specifies IP-derived routes for Intermediate System-to-Intermediate System (IS-IS).

Configuration Commands to Control Routing Updates in Redistribution

As explained in the section “Controlling Routing Updates During Redistribution,” it is necessary to control the flow of updates between the routing protocol as well as throughout the autonomous system. The following sections consider the implementation of passive interfaces, static routes, and default routes.

Configuring the Passive Interface

The passive interface is used for routing protocols that send updates through every interface with an address that is included in the **network** command. If the routing protocol is not running on the next-hop router, it is a waste of time to send updates out of the interface.

The command reduces spending limited resources without compromising the integrity of the router. The router processes all routes received on an interface.

The command syntax to configure a passive interface, where *type* and *number* indicate the interface to be made, is as follows:

```
Router(config-router)#passive-interface type number
```

Configuring Static Routes

The following shows the syntax for configuring the static route:

```
Router(config)#ip route prefix mask {address|interface} [distance] [tag tag] [permanent]
```

This defines the path by stating the next-hop router to which to send the traffic. This configuration can be used only if the network address for the next-hop router is in the routing table. If the static route needs to be advertised to other routers, it should be redistributed.

In some versions of the IOS software, this route is automatically redistributed. It is viewed as a connected network, as long as the output interface is referred in the static route instead of an IP address.

Table 17-8 explains the options available in the static route command.

Table 17-8 *Explanation of the IP Route Options*

Command	Description
<i>prefix</i>	The route prefix for the destination.
<i>mask</i>	The prefix mask for the destination.
<i>address</i>	The IP address of the next-hop router that can be used to reach that network.

(continues)

Table 17-8 Explanation of the IP Route Options (Continued)

Command	Description
<i>Interface</i>	The network interface to use to get to the destination network.
<i>distance</i>	Optional administrative distance to assign to this route. (Recall that administrative distance refers to how believable the routing protocol is.)
tag tag	Optional value that can be used as a match value in route maps.
<i>Permanent</i>	Specification that the route will not be removed even if the interface associated with the route goes down.

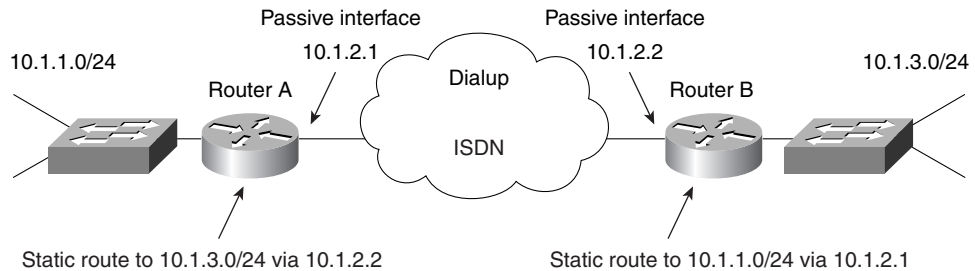
Use static routes pointing to the outgoing interface only on point-to-point interfaces. Static routes configured on multipoint or multiaccess interfaces need a next-hop address. On point-to-point interfaces, the information is sent to the only other device on the network.

In Example 17-4 (and the corresponding Figure 17-8), the use of a static route and the **passive-interface** command is illustrated. Additional configuration is included to place the commands in context; the commands relevant to this section are placed in bold.

Example 17-4 The Use of Static Routing and Passive Interfaces

```
Router(config)#Router A
RouterA(config)#username RouterB password Shhh
RouterA(config)#dialer-list 1 protocol ip permit
RouterA(config)#interface bri 0
RouterA(config-if)#encapsulation ppp
RouterA(config-if)#ip addr 10.1.2.1 255.255.255.0
RouterA(config-if)#ppp authentication chap
RouterA(config-if)#dialer map ip 10.1.2.2 broadcast name RouterB 1222555222201
RouterA(config-if)#dialer-group 1
RouterA(config)#interface ethernet 0
RouterA(config-if)#ip address 10.1.1.1 255.255.255.0
RouterA(config)#ip route 10.1.3.0 255.255.255.0 10.1.2.2
RouterA(config)#router eigrp 1
RouterA(config-router)#network 10.0.0.0
RouterA(config-router)#passive-interface Bri0
```

In this example, the link between Routers A and B is raised when Router A sees interesting traffic try to exit the serial interface. *Interesting traffic* is traffic that is permitted in a preconfigured access list. In this example, all IP traffic is considered interesting. This example is perfectly valid for occasional access, except for the few additional ISDN parameters that need to be added. Figure 17-8 illustrates the use of both static routes and passive interfaces.

Figure 17-8 *The Use of Static Routes Across a Dialup Link*

Note: No routing protocol is running across the ISDN cloud.

In this example, you see that EIGRP updates do not flow across the dialup line because the interface pointing into the ISDN cloud is configured as a passive interface. The same configuration has been applied to Router B so that no updates raise the WAN link.

Neither Router A nor Router B knows of the networks on the other side of the WAN, so static routes must be configured.

Configuring Default Routes

In larger networks, there might be many static routes to be configured. Not only is this a chore for the administrator, but it also requires vigilance so that changes in the routing table can be reconfigured. It might be that turning on a routing protocol is advised, or alternatively, you can configure a specialized static route, called a *static default route*.

The following is a static default route that will generate a default route on the router configured:

```
Router(config)#ip route 0.0.0.0 0.0.0.0 s0
```

NOTE The different routing protocols treat these default route commands differently when redistributing them into the routing protocol. Reference the Cisco documentation set for detailed explanations.

The default routes are propagated through the network dynamically or can be configured into the individual routers.

If a router has a directly connected interface onto the specified default network, the dynamic routing protocols running on that router will generate or source a default route. In the case of RIP, it will

advertise the pseudonetwork 0.0.0.0. In the case of IGRP, the network itself is advertised and flagged as an exterior route.

When default information is being passed along through a dynamic routing protocol, no further configuration is required. In the case of RIP, there can be only one default route, network 0.0.0.0. However, in the case of IGRP, several networks can offer default routes, although only one is used.

If the router is not directly connected to the default network but does have a route to it, it is considered as a candidate default path. To configure a default route, use the following syntax:

```
Router(config)#ip default-network network-number
```

This command will generate a default route to be sent in updates. It does not generate a default network on the router that was configured, and it will only generate a default route if the route used is directly connected. When there are multiple default routes in the routing table, the route candidates are examined. As you would expect, the best default path is selected based on administrative distance and metric. The gateway to the best default path then becomes the gateway of last resort for the router, which is another term for default router. You can display the gateway of last resort with this command:

```
Router#show ip route
```

The default route will appear in the routing table marked as a static route with S*. The gateway of last resort will be set to this network.

Redistribution Examples

The following examples are case studies that pull together the concepts you learned about redistribution. Redistribution involves complex design and configuration considerations. Therefore, it is best to see the various problems and solutions illustrated in context.

This section presents three examples:

- Route redistribution without redundant paths between different routing protocols.
- Route redistribution with redundant paths between different routing protocols. The example also covers resolving the path selection problems that result in redistributed networks.
- The use of a default network in a redistributed environment.

Example 1: Route Redistribution Without Redundant Paths

Refer to Figure 17-9 for this example of route redistribution without redundant paths between different routing protocols.

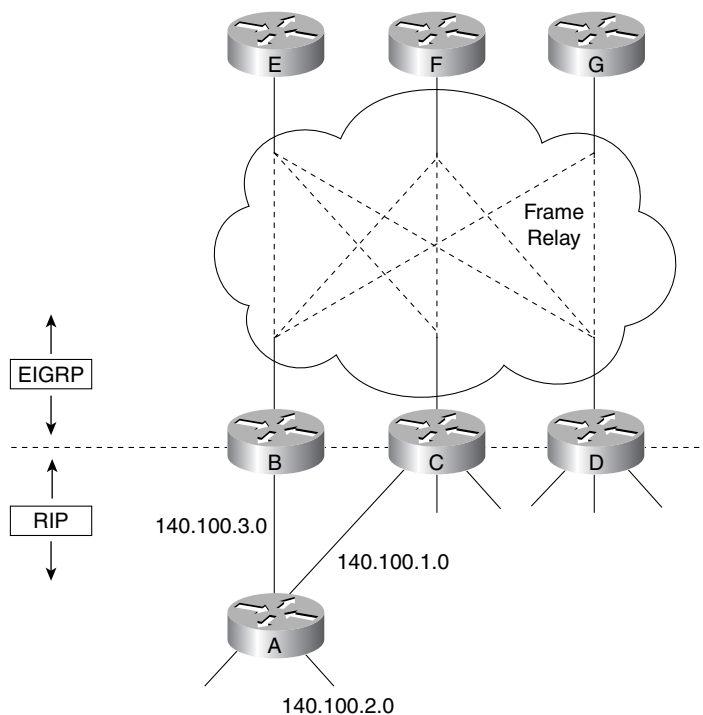
In the future, the company might want to add redundancy by meshing the Frame Relay cloud and consolidating the three core routers into one large router. Currently, the company has a simple and low-cost solution using existing equipment.

Example 2: Route Redistribution with Redundant Paths

Refer to Figure 17-10 for this example, which covers route redistribution with redundant paths between different routing protocols and resolving path selection problems that result in redistributed networks.

In Figure 17-10, Router A is connected to networks 140.100.1.0, 140.100.2.0, and 140.100.3.0. Using RIP, network 140.100.1.0 is advertised to Router B, 140.100.3.0 is advertised to Router C, and network 140.100.2.0 is advertised to both Routers A and B.

Figure 17-10 *Choosing the Optimal Path, Through Administrative Distance, When Redistribution Is Using Redundant Paths*



The routing table of Router A will show the information presented in Table 17-9.

Table 17-9 Router A Routing Table Information

Routing Protocol	Network/Subnet	Next Logical Hop	Metric
Connected	140.100.1.0/24	Connected E0	0
Connected	140.100.2.0/24	Connected E1	0
Connected	140.100.3.0/24	Connected E2	0
RIP	10.10.10.8/30	140.100.3.2	1 hop
RIP	10.10.10.12/30	140.100.3.2	1 hop
RIP	10.10.10.16/30	140.100.3.2	1 hop
RIP	10.10.10.20/30	140.100.3.2	1 hop
RIP	10.10.10.24/30	140.100.3.2	1 hop
RIP	10.10.10.28/30	140.100.3.2	1 hop
RIP	10.10.10.32/30	140.100.3.2	1 hop
RIP	193.144.6.0/24	140.100.3.2	1 hop
RIP	193.144.7.0/24	140.100.3.2	1 hop
RIP	193.144.8.0/24	140.100.3.2	1 hop

The routing table of Router B will show the information presented in Table 17-10.

Table 17-10 Router B Routing Table Information

Routing Protocol	Network/Subnet	Next Logical Hop	Metric
RIP	140.100.1.0/24	140.100.3.1	1 hop
RIP	140.100.2.0/24	140.100.3.1	1 hop
Connected	140.100.3.0/24	Connected E0	0
Connected	10.10.10.8/30	Connected S0	0
Connected	10.10.10.12/30	Connected S0	0
Connected	10.10.10.16/30	Connected S0	0
EIGRP	10.10.10.20/30	10.10.10.9	2221056
EIGRP	10.10.10.24/30	10.10.10.9	2221056

(continues)

Table 17-10 Router B Routing Table Information (Continued)

Routing Protocol	Network/Subnet	Next Logical Hop	Metric
EIGRP	10.10.10.28/30	10.10.10.13	2221056
EIGRP	10.10.10.32/30	10.10.10.13	2221056
EIGRP	193.144.6.0/24	10.10.10.9	2221056
EIGRP	193.144.7.0/24	10.10.10.13	2221056
EIGRP	193.144.8.0/24	10.10.10.17	2221056

Note that the routing table for Router A sees all the subnets for network 10.0.0.0 with a mark of 255.255.255.252 or /30. However, because RIP does not pass the network mark in updates and Router A is not connected to network 10.0.0.0, a static route must have been configured so that Router A can see the /30 mask.

The routing table sees all the paths as unique, so it is clear which paths are accessible through RIP or EIGRP. Even after redistribution, the routing table will not change; the confusion occurs after the propagation of the EIGRP updates through the network.

The EIGRP updates will be sent to all the routers in the domain, and Routers E, F, and G will have no confusion. Depending on the timing of the updates and convergence, however, Router C might become confused. Routers E, F, and G will have sent information on how to get to the networks 140.100.1.0 and 140.100.2.0. Router C will also receive information from Router A. Sending the data traffic to Router A is obviously the optimum path; however, because EIGRP has a significantly better administrative distance, the EIGRP route will be placed in the routing table as having the best path. On the assumption that the Frame Relay PVCs all have the same bandwidth, the routing table will see all three paths and distribute the traffic evenly among them.

Example 17-5 shows how to configure Routers B, C, and D to change the administrative distance to favor RIP for the LANs within its domain. The networks 140.100.1.0 and 140.100.2.0 are given an administrative distance of 200 in accordance with the access list. This ensures that the RIP path will be favored if it is available.

Example 17-5 Changing the Administrative Distance to Favor RIP

```
Router(config)#router rip
Router(config-router)#network 140.100.0.0
Router(config-router)#passive interface S0.1
Router(config-router)#redistribute eigrp 100 metric 3
Router(config)#router eigrp 100
Router(config-router)#network 140.100.0.0
Router(config-router)#passive interface E0
```

Example 17-5 *Changing the Administrative Distance to Favor RIP (Continued)*

```

Router(config-router)#default-metric 10000 100 255 1 1500
Router(config-router)#distance 200 0.0.0.0 255.255.255.255 3
Router(config)#access-list 3 permit 140.100.1.0
Router(config)#access-list 3 permit 140.100.2.0

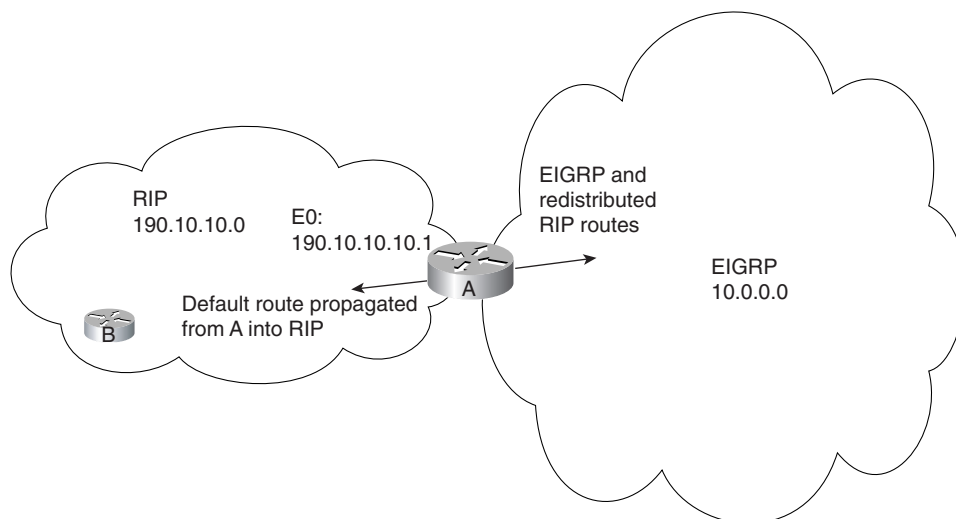
```

The **distance** command sets the administrative distance for the EIGRP 100 process. It changes the distance from 90 to 200, which makes the routes that RIP offers more favorable because RIP has an administrative distance of 120. The use of 0.0.0.0 with a wildcard mask of 255.255.255.255 is just as a placeholder. It indicates that although the command allows for a network to be specified so that the administrative distance can be applied selectively to that network, in this configuration, no network has been selected. The command has been applied to all networks. You do want the administrative distance to be altered on two networks, however. This granularity cannot be stated in the **distance** command; therefore, an access list is used. In the example, the number 3 at the end of the command line points to the access list that carries that number as an identifier. The access list, by permitting networks 140.100.1.0 and 140.100.2.0, is identifying the networks to which the **distance** command is to be applied.

Example 3: A Default Network in a Redistributed Environment

The use of the default network simplifies the configuration of a redistributed network by allowing the redistribution to be one-way. This significantly reduces the possibility of feedback of networks into the originating domain. The configuration for this example is inset within Figure 17-11 because the configuration of more than one router is shown.

Figure 17-11 *The Use of a Default Network in a Redistributed Network to Resolve Problems with Path Selection*



In this design, every router and, therefore, workstation within the RIP domain sees its own internal networks, but all other networks are accessed via the default route. Router B's configuration is shown in Example 17-6.

Example 17-6 *Router B Configuration*

```
RouterB(config)#router rip
RouterB(config-router)#network 190.10.10.0
```

Router A redistributes between RIP and EIGRP and acts as an ABR in OSPF, with the RIP domain acting as a stub network. The default route is configured as a static route on Router A, redistributed into RIP, and propagated throughout the RIP domain. The internal RIP-only routers must be configured to accept a default route with a destination network because it is only reachable via a route default.

The configuration for Router A is shown in Example 17-7.

Example 17-7 *Router A Configuration*

```
RouterA(config)#router eigrp 100
RouterA(config-router)#redistribute rip
RouterA(config-router)#default-metric 10000 100 255 1 1500
RouterA(config-router)#network 10.0.0.0

RouterA(config)#router rip
RouterA(config-router)#network 190.10.10.0
RouterA(config-router)#redistribute static

RouterA(config)#ip route 0.0.0.0 0.0.0.0 10.10.10.1
```

The redistribution on Router A can now be one-way. EIGRP needs to know all the networks in the RIP domain, but RIP, when configured with a default route, needs no understanding of the outside world. The RIP domain works in a similar fashion as a stub network in OSPF.

Controlling Routing Updates with Filtering

Despite all the mechanisms for controlling and reducing the routing updates on your network, it is sometimes necessary to wield greater and more flexible power. This comes in the form of access lists, which when applied to routing updates are referred to as *distribute lists*.

The logic used in the distribute list is similar to that of an access list. The process is listed in the following text:

1. The router receives a routing update or is about to send a routing update about one or more networks.
2. The router looks at the appropriate interface involved with the action to check for filtering.
3. The router determines whether a filter is associated with the interface.
4. If a filter is present, the router examines the access list to see if there is a match on any of the networks in the routing update.
5. If there is no filter on the interface, the routing update is sent directly to the routing process as normal.
6. If there is a filter, the route entry is processed according to the distribute list: advertise the route if matched by a **permit** statement or do not advertise if it is matched by a **deny** statement.
7. If no match is found in the distribute list, the implicit **deny any** at the end of the access list will cause the update to be dropped.

Routing updates can be filtered for any routing protocol by defining an access list and applying it to a specific routing protocol. There are some limitations to distribute lists when applied to OSPF. For example, the inbound list prevents routes entering the routing table but does not prevent link-state packets from being propagated.

When creating a routing filter or distribute list, the following steps should be taken:

- Write out in longhand what you are trying to achieve.
- Identify the network addresses to be filtered, and create an access list. Permit the networks you want to have advertised.
- Determine whether you are filtering routing updates coming into the router or updates to be propagated to other routers.
- Assign the access list using the **distribute-list** command.

Use the following command syntax to configure the distribute list to filter incoming updates:

```
Router(config-router)#distribute-list {access-list-number | name} in [type number]
```

Table 17-11 explains the options of this command.

Table 17-11 *Explanation of the **distribute-list in** Command Options*

Command	Description
<i>access-list-number name</i>	Gives the standard access list number or name
in	Applies the access list to incoming routing updates
<i>type number</i>	Gives the optional interface type and number from which updates will be filtered

Use the following command syntax to configure the distribute list to filter outgoing updates:

```
Router(config-router)#distribute-list {access-list-number | name} out
[interface-name | routing-process | autonomous-system-number]
```

Table 17-12 explains the options of this command.

Table 17-12 *Explanation of the **distribute-list out** Command Options*

Command	Description
<i>access-list-number name</i>	Gives the standard access list number or name
out	Applies the access list to outgoing routing updates
<i>interface-name</i>	Gives the optional interface name out of which updates will be filtered
<i>routing-process</i>	Gives the optional name of the routing process, or the keyword static or connected , from which updates will be filtered
<i>autonomous-system-number</i>	Gives the optional autonomous system number of routing process

Verifying, Maintaining, and Troubleshooting the Implementation of Redistribution and Filtering

The key to maintaining and troubleshooting the redistribution within your network is to have a clear understanding of the network topology from both a physical and a logical perspective. The traffic flows—the peaks and lows in the traffic volume—are also important in truly understanding the connectivity issues within the network. From this vantage point, it is possible to interpret the output presented by the various tools available.

Most of the appropriate commands in tracking redistribution problems are ones that have been examined earlier. They include the following:

- **show ip protocol**
- **show ip route**
- **show ip route** *routing-protocol*
- **show ip eigrp neighbors**
- **show ip ospf database**

In addition to these commands, the **trace** and extended **ping** commands are also very useful.

The trace Command

The **trace** command is invoked from user mode, whereas the extended **trace** is only available from the exec privileged level. This shows the routers that a packet has passed through to reach its destination.

The extended **trace** test is called by entering the command without any destination. This results in the utility asking a series of questions, allowing you to change the defaults.

The Extended ping Command

To check host reachability and network connectivity, use the **ping** privileged exec command. The extended **ping** utility is called by entering the command without any destination. This results in the utility asking a series of questions, allowing you to change the defaults.

Using trace and Extended ping

You do not use **trace** to determine the path taken, but rather to identify where there is a problem in the network. Where the **trace** utility fails indicates a good starting point for troubleshooting a complex network.

The **trace** command is not very useful in reflecting the routing path because path changes are not shown. The extended **ping** command, however, is very useful because it announces every interface that it traverses if the record option is selected. The limitation is the maximum hops that it can report, which is nine.

It is also possible to specify a source address in the **trace** or **ping** commands (as long as it is an interface on the router). This can be useful for testing certain types of access lists, route maps, and so on. Otherwise, the route will choose the source address of its own interface closest to the destination. It is also useful for testing network reachability from the far end.

These commands are generic to TCP/IP troubleshooting.

Foundation Summary

The “Foundation Summary” section of each chapter lists the most important facts from the chapter. Although this section does not list every fact from the chapter that will be on your exam, a well-prepared candidate should, at a minimum, know all the details in each “Foundation Summary” before going to take the exam.

Various methods enable you to control the routing information sent between routers. These methods include the following:

- **Passive interfaces** — An interface that does not participate in the routing process. In RIP and IGRP, the process listens but will not send updates. In OSPF and EIGRP, the process neither listens nor sends updates because no neighbor relationship can form.

The interfaces that participate in the interior routing process are controlled by the interface configuration. During configuration, the routing process is instructed via the **network** command on which interfaces to use. Because most protocols express the networks at the major boundary, interfaces that have no reason to send this protocol’s updates propagate the data across the network. This is not only a waste of bandwidth but, in many cases, can also lead to confusion.

- **Default route** — A route used if there is no entry in the routing table for the destination network. If the lookup finds no entry for the desired network and no default network is configured, the packet is dropped.

If the routing process is denied the right to send updates, the downstream routers will have a limited understanding of the network. To resolve this, use default routes. Default routes reduce overhead, add simplicity, and can remove loops.

- **Static routes** — A route that is manually configured. It takes precedence over dynamic routes learned via a routing process.

If no routing process is configured, static routes might be configured to populate the routing table. This is not practical in a large network because the table cannot learn of changes in the network topology dynamically. In small environments or for stub networks, however, this is an excellent solution.

- **The null interface** — An imaginary interface that is defined as the next logical hop in a static route. All traffic destined for the remote network is carefully routed into a black hole.

This can be used in a similar way as the passive interface, but it allows for greater granularity in the denied routes.

It is also used to feed routes into another routing protocol. It allows another mask to be set and, therefore, is useful when redistribution occurs between a routing protocol that uses VLSM and one that does not. In this way, it aggregates routes as shown in the previous chapter.

- **Distribute lists** — Access lists applied to the routing process, determining which networks will be accepted into the routing table or sent in updates.

When communicating to another routing process, it is important to control the information sent into the other process. This control is for security, overhead, and management reasons. Access lists afford the greatest control for determining the traffic flow in the network.

- **Route maps** — Complex access lists permitting conditional programming. If a packet or route matches the criteria defined in a **match** statement, changes defined in the **set** command are performed on the packet or route in question.

Table 17-13 shows the requirements of automatic redistribution between routing protocols.

Table 17-13 *Automatic Redistribution Between Routing Protocols*

Routing Protocol	Redistribution Policy
Static	Requires manual redistribution into other routing protocols.
Connected	Unless included in the network command for the routing process, requires manual redistribution.
RIP	Requires manual redistribution.
IGRP	Will automatically redistribute between IGRP and EIGRP if the autonomous system number is the same. Otherwise, processes with different IGRP autonomous system numbers, or IGRP and EIGRP processes with different autonomous system numbers, require manual redistribution.
EIGRP	Will automatically redistribute between IGRP and EIGRP if the autonomous system number is the same. Otherwise, processes with different EIGRP autonomous system numbers, or IGRP and EIGRP processes with different autonomous system numbers, require manual redistribution. EIGRP for AppleTalk will automatically redistribute between EIGRP and RTMP. EIGRP for IPX will automatically redistribute between EIGRP and IPX RIP/ SAP; in later versions, NLSP can be manually redistributed.
OSPF	Requires manual redistribution between different OSPF process IDs and routing protocols.

The following list explains the logic used in a distribute list:

1. The router receives a routing update or is about to send a routing update about one or more networks.
2. The router looks at the appropriate interface involved with the action to check for filtering.
3. The router determines whether a filter is associated with the interface.
4. If a filter is present, the router examines the access list to see if there is a match on any of the networks in the routing update.
5. If there is no filter on the interface, the routing update is sent directly to the routing process, as normal.
6. If there is a filter, the route entry is processed according to the distribute list: advertise the route if matched by a **permit** statement or do not advertise if it is matched by a **deny** statement.
7. If no match is found in the distribute list, the implicit **deny any** at the end of the access list will cause the update to be dropped.

Figure 17-12 shows the decisions that are made when a distribute list is applied to an incoming routing update, and Figure 17-13 shows the decisions made for an outgoing routing update. The logic used in a distribute list is similar to an access list and is summarized in the following chart. Refer to the section “Controlling Routing Updates with Filtering” for a detailed discussion of this subject.

Figure 17-12 *Distribute List Logic on an Incoming Update*

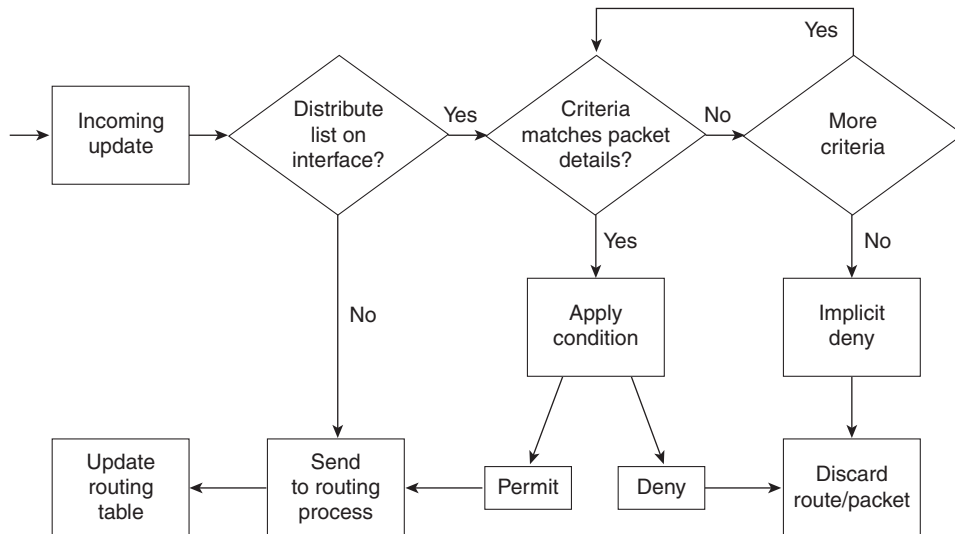
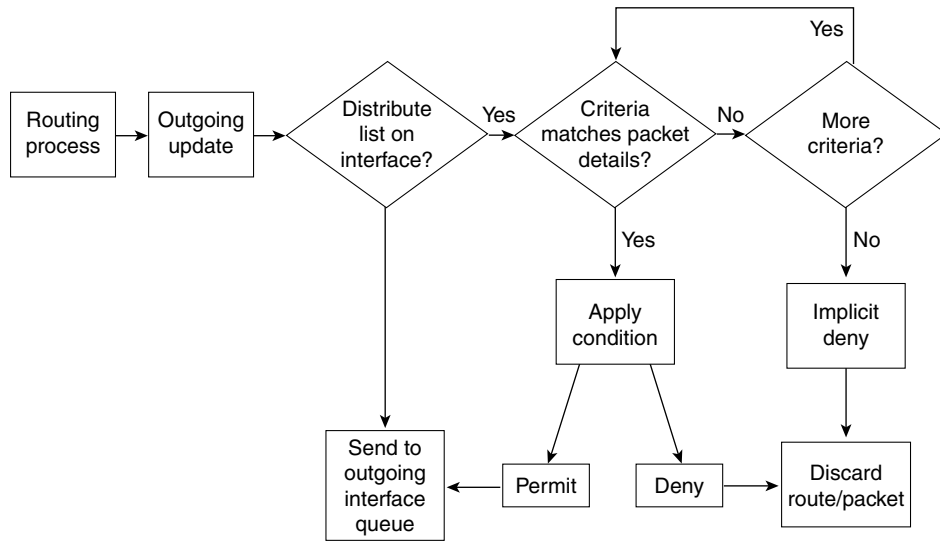


Figure 17-13 *Distribute List Logic on an Outgoing Update*

Q&A

As mentioned in the introduction, “All About the CCNP, CCDP, and CCIP Certifications,” you have two choices for review questions. The questions that follow next give you a bigger challenge than the exam itself by using an open-ended question format. By reviewing now with this more difficult question format, you can exercise your memory better and prove your conceptual and factual knowledge of this chapter. The answers to these questions are found in Appendix A.

For more practice with examlike question formats, including questions using a router simulator and multichoice questions, use the exam engine on the CD-ROM.

1. State two of the methods that Cisco recommends for controlling routing protocol traffic.
2. What is the default administrative distance for RIP?
3. State two instances when you do not want routing information propagated.
4. In what instances will EIGRP automatically redistribute?
5. Which command is used to view the administrative distance of a route in the routing table?
6. When is redistribution required?
7. Why does Cisco recommend that you not overlap routing protocols?
8. Why would you want to prevent routing updates across an on-demand WAN link?
9. What is the metric used for in a routing protocol?
10. Give two reasons for using multiple routing protocols.
11. In a very large environment, the various domains might have different requirements, making a single solution inefficient. A clear example is the case of a large multinational corporation, where EIGRP is the protocol used at the access and distribution layers, but BGP is the protocol connecting the core. When implementing redistribution, state one possible problem that you might experience, and explain why it is a problem.
12. Which has a lower administrative distance, IGRP or OSPF?
13. What command is used to configure an outbound route filter?
14. What is a passive interface?
15. What is the purpose of administrative distance?
16. What is the concern of redistributing into a redundant network?

17. What is a default network?
18. Why is it necessary to configure a default metric when redistributing between routing protocols?
19. Which command is used to modify the administrative distance of a route?
20. What is the difference in processing for an inbound and an outbound route filter?

Scenario

The following scenarios and questions are designed to draw together the content of the chapter and to exercise your understanding of the concepts. There is not necessarily a right answer. The thought process and practice in manipulating the concepts are the goals of this section. The answers to the scenario questions are found at the end of this chapter.

Scenario 17-1

Duddleduddle is a large hospital with several sites in the city. Although the sites connect to a centralized patient and administration database, the hospital has fought for local autonomy based on the specialization of the site and the fact that it is its own business unit. An IT group manages the central administration and oversees the other sites. The chief information officer (CIO) who ran this group and the overall network has left because of political wrangling. The new CIO, recently appointed, is attempting to sort out the mess.

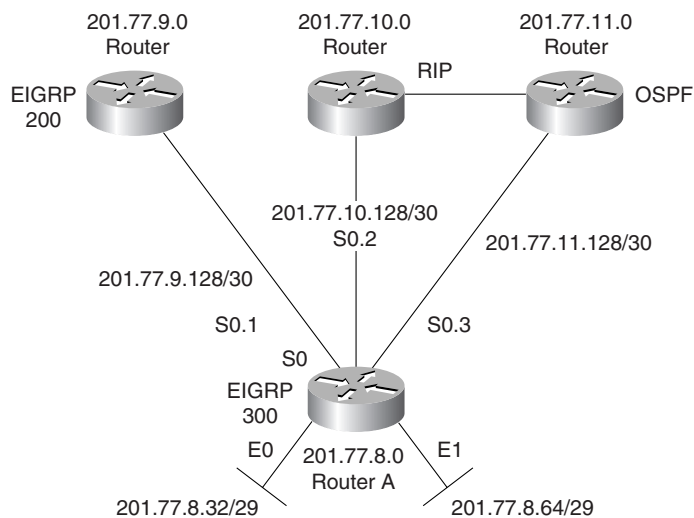
This new CIO has the agreement of the other hospital sites that there should be one routing protocol, as opposed to the four that are currently running. In turn, he has agreed to implement filtering to improve the network performance, grant some basic security, and indulge some turf wars.

The first step to creating a single routing protocol network is to redistribute the protocols so that the network can see all the available routes. Unfortunately, the routing protocols are aware of multiple path destinations. Therefore, the implementation must be done not only with consideration to preventing routing loops, but also with optimal path selection.

Figure 17-14 shows the network topology for the hospital Duddleduddle.

Using the figure as reference, complete the following exercises.

1. Issue the configuration commands for the RIP network to be redistributed on Router A into EIGRP.
2. On Router A, ensure that the interfaces running EIGRP do not have RIP updates generated through them or that the RIP interfaces do not have EIGRP updates running through them.
3. The site running IGRP and the site running EIGRP are running different autonomous system numbers. How would you implement a transition to both sites running EIGRP using the same autonomous system number?

Figure 17-14 *Topology for the Scenario 17-1 Network*

4. The OSPF redistribution into RIP has been implemented, but users are complaining about delays. State the first step that you would take to verify the configuration.
5. The CIO has been asked to submit a transition plan to the board of trustees that includes a reasoned explanation for the need for redistribution. What should it look like?

Scenario Answers

The answers provided in this section are not necessarily the only possible answers to the questions. The questions are designed to test your knowledge and to give practical exercise in certain key areas. This section is intended to test and to exercise skills and concepts detailed in the body of this chapter.

If your answer is different, ask yourself whether it follows the tenets explained in the answers provided. Your answer is correct not if it matches the solution provided in the book, but rather if it has included the principles of design laid out in the chapter.

In this way, the testing provided in these scenarios is deeper: It examines not only your knowledge, but also your understanding and ability to apply that knowledge to problems.

If you do not get the correct answer, refer back to the text and review the subject tested. Be certain also to review your notes on the question to ensure that you understand the principles of the subject.

Scenario 17-1 Answers

1. *Issue the configuration commands for the RIP network to be redistributed on Router A into EIGRP.*

The commands are as follows:

```
Router(config)#router eigrp 300
Router(config-router)#network 210.77.8.0
Router(config-router)#redistribute rip
Router(config-router)#default-metric 10000 100 255 1 1500
```

2. *On Router A, ensure that the interfaces running EIGRP do not have RIP updates generated through them or that the RIP interfaces do not have EIGRP updates running through them.*

Strictly speaking, the **passive-interface** command on s0.2 is not required because the interface address is different from that stated in the **network** command. The same is true for RIP. To ensure this:

```
Router(config)#router eigrp 300
Router(config-router)#network 210.77.8.0
Router(config-router)#redistribute rip
Router(config-router)#default-metric 10000 100 255 1 1500
Router(config-router)#passive-interface s0.2
Router(config)#router rip
Router(config-router)#network 201.77.10.0
Router(config-router)#passive-interface s0.1
Router(config-router)#passive-interface s0.3
Router(config-router)#passive-interface e0
Router(config-router)#passive-interface e1
```


3. *The site running IGRP and the site running EIGRP are running different autonomous system numbers. How would you implement a transition to both sites running EIGRP using the same autonomous system number?*

There are several ways to transition from different autonomous systems to one autonomous system so that IGRP and EIGRP automatically redistribute. The methods include the following:

- Configuring redistribution at both sites, and in a controlled manner during downtime, switching all the routers in the IGRP site to the same autonomous system as that of EIGRP. Because the prevailing routing protocol is to be EIGRP, it makes sense that IGRP is the protocol to have its autonomous system number changed.
 - Another approach, and one favored by many, is to configure EIGRP with the same autonomous system number on all routers at the IGRP site. As part of the configuration, increase the administrative distance of EIGRP to be 200 so that none of the routes is acceptable to the routing process. Then, during downtime on the systems, cut over to the EIGRP process by changing its administrative distance back to the default of 90. This can be done by simply adding the word *no* in front of the existing command. The beauty of this plan is that everything can be put in place before the cutover; if problems are experienced, it is equally easy to reverse the command to return to the IGRP configuration.
4. *The OSPF redistribution into RIP has been implemented, but users are complaining about delays. State the first step that you would take to verify the configuration.*

The first step is to issue the following commands, perhaps on both the OSPF and RIP routers:

- **show ip route** : To ensure that each routing process sees the appropriate paths. A routing loop might be visible here.
- **extended ping** : To see the path that is taken to the remote locations. A routing loop could be detected.
- **show ip protocols** : To see how RIP and OSPF are being redistributed, what the default metrics are, and whether there are any distribute lists impeding the flow of updates.
- **show ip ospf database** : To ensure that all the routes are in place. Again, errors leading to a routing loop could be detected here.
- **show ip ospf neighbor** : To ensure that OSPF can still see the adjacent routers.

5. *The CIO has been asked to submit a transition plan to the board of trustees that includes a reasoned explanation for the need for redistribution. What should it look like?*

It should include the following:

- The transition would happen at the main site where the centralized records and databases are maintained. This is because this site must be the most stable because it serves the other sites.
- The next step would be to review the addressing scheme to ensure that it was hierarchical and could support summarization and VLSM.
- The redistribution between IGRP and EIGRP is the easiest to effect and should be performed in accordance to the answer in question 3.
- Because the non-Cisco Systems equipment for RIP and OSPF machines might not support EIGRP, a careful assessment should be done, and plans should be made to upgrade the equipment as necessary. However, the network administrator should be reminded that it is not necessary for hosts to run RIP. If hosts run RIP, it can lead to a very unstable network. Therefore, all hosts should be configured with a default gateway address.
- Configure redistribution in the network to ensure the full connectivity throughout the campuses.
- When redistribution is in place, centralization of resources and maintenance of the data and network can be implemented, granting a full exchange of information throughout the hospital to harness the power of the information available.



This chapter covers the following topics, which you need to understand to pass the CCNP/CCDP/CCIP BSCI exam:

- Understanding route maps
- Understanding policy-based routing
- The operation of route maps and policy-based routing
- Configuring route maps for policy-based routing
- Configuring fast switching with policy-based routing
- Configuring route maps for redistribution
- Monitoring the configuration of route maps, policy-based routing, and redistribution

Controlling Network Traffic with Route Maps and Policy-Based Routing

The topics in this chapter deal with controlling both routed and routing traffic with route maps, which are more sophisticated than access lists. This is an advanced topic that deals with programming the router to match criteria against assigned lists and to perform tasks based on the result of the match.

The chapter deals with why route maps are needed and how they work. This chapter also provides the configuration syntax with working examples.

Route maps are rather intimidating if you are not familiar with access lists. Access lists are dealt with in depth in the CCNA course materials. The books *CCNA Self-Study: Interconnecting Cisco Network Devices (ICND)* and the *CCNA ICND Exam Certification Guide*, both from Cisco Press, deal with these subjects in more depth.

“Do I Know This Already?” Quiz

The purpose of the “Do I Know This Already?” quiz is to help you to decide what parts of this chapter to use. If you already intend to read the entire chapter, you do not necessarily need to answer these questions now.

The 17-question quiz, derived from the major sections in the “Foundation Topics” portion of the chapter, helps you to determine how to spend your limited study time.

Table 18-1 outlines the major topics discussed in this chapter and the “Do I Know This Already?” quiz questions that correspond to those topics.

Table 18-1 “Do I Know This Already?” Foundation Topics Section-to-Question Mapping

Foundation Topics Section	Questions Covered in This Section
Understanding Route Maps	1–5
Understanding Policy-Based Routing	6–8
The Operation of Route Maps and Policy-Based Routing	9–10
Configuring Route Maps for Policy-Based Routing	11–12
Configuring Fast Switching with Policy-Based Routing	13–14
Configuring Route Maps for Redistribution	15–16
Monitoring the Configuration of Route Maps, Policy-Based Routing, and Redistribution	17

CAUTION The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark this question wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

1. In what ways are route maps more sophisticated than access lists?
 - a. Access lists can change the destination address of the outbound data packet.
 - b. Route maps take less CPU because of streamlined processing.
 - c. Access lists either exclude or include, whereas route maps can change the characteristics of the path.
 - d. Access lists can only be applied on inbound updates.
2. Route maps can be used for which of the following applications?
 - a. NAT
 - b. BGP
 - c. Summarization
 - d. Redistribution

3. Which of the following best describe a **match** statement?
 - a. The means by which a route is selected
 - b. A list of selection criteria
 - c. The method of discarding unwanted packets
 - d. A list of network prefixes
4. Which of the following best describe a **set** statement?
 - a. The method used to determine the best metric
 - b. The means of choosing the next hop
 - c. A list of conditions to apply to chosen routes
 - d. The means of changing routes or packets that are matched
5. Many **match** statements can be used in a route map. How many **match** statements must be matched for the **set** to be applied?
 - a. At least one
 - b. All of the criteria
 - c. None of them
 - d. At least 50 percent of the criteria
6. What are some of the benefits of policy-based routing?
 - a. The ability to link extended access lists
 - b. Easy administration
 - c. QoS
 - d. Load balancing
7. What is the relationship between route maps and policy-based routing?
 - a. Route maps use policy-based routing.
 - b. Route maps and policy-based routing are interchangeable terms.
 - c. Policy-based routing uses route maps.
 - d. Policy-based routing works with access lists, whereas route maps use a programming language.

8. Policy-based routing is applied to what type of traffic?
 - a. Inbound and routed traffic dependent on the configuration
 - b. Incoming packets
 - c. The routing process
 - d. Transiting traffic
9. If no match is made in a policy-based routing list of criteria, what action is taken?
 - a. The packet is sent to null interface 0.
 - b. The packet is dropped and no ICMP packet is sent to the source.
 - c. The packet is dropped and an ICMP packet is sent to the source.
 - d. The packet is sent to the routing process.
10. What additional configuration is required to ensure that packets are dropped when matched to a deny statement?
 - a. The **set** command at the end of the route map should be a static route to null 0.
 - b. The deny statement should be set to 0.0.0.0 0.0.0.0.
 - c. The last **set** command should be configured to the no-forward parameter.
 - d. No extra configuration is required, because packets that are denied are automatically dropped.
11. What command is used to match a packet based on its size?
 - a. **match length**
 - b. **match size**
 - c. **match MTU**
 - d. **match ip length**
12. Which **set** command is used only when there is no route found in the routing table?
 - a. **set ip next hop**
 - b. **set default next hop**
 - c. **set ip default next-hop**
 - d. **set default gateway**

13. What command is used to configure fast switching for route maps?
 - a. No command is required, because fast switching is on by default
 - b. Router(config-route-map)#**set fast-switch on**
 - c. Router(config-if)#**ip route-cache policy**
 - d. router#**enable fast-switching**
14. When using route maps for redistribution when there is a match and the deny statement is configured, what action will be taken by the route map?
 - a. The packet is dropped.
 - b. The route is not redistributed.
 - c. An ICMP packet is sent to the sender.
 - d. The packet is sent to the normal routing process.
15. How are the services of the route map for redistribution called?
 - a. The **route map** command
 - b. Under the incoming interface
 - c. The **redistribution** command
 - d. As a global configuration command
16. Which command is useful for validating the path to the destination with a specified packet size?
 - a. Extended **ping**
 - b. **show ip protocol**
 - c. **show ip route**
 - d. **show redistribution**
17. Which command is used to show the configured route maps?
 - a. **show ip route map**
 - b. **show route-map**
 - c. **show ip map**
 - d. **show ip policy**

The answers to this quiz are found in Appendix A, “Answers to Chapter ‘Do I Know This Already?’ Quizzes and Q&A Sections.” The suggested choices for your next step are as follows:

- **9 or less overall score** —Read the entire chapter. This includes the “Foundation Topics” and “Foundation Summary” sections, the “Q&A” section, and the “Scenarios” at the end of the chapter.
- **10–14 overall score** —Begin with the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. If you have trouble with these exercises, read the appropriate sections in “Foundation Topics.”
- **15 or more overall score** —If you want more review on these topics, skip to the “Foundation Summary” section, and then go to the “Q&A” section and the “Scenarios” at the end of the chapter. Otherwise, move to the next chapter.

Foundation Topics

Understanding Route Maps

Route maps are the means by which sophisticated “if/then logic” can be applied to a router. Route maps are the programming tools that are used to control redistribution, to implement policy-based routing, to control NAT translation, and to implement BGP policy.

You can use route maps for the following purposes:

- **To control redistribution** —Route maps allow a higher level of sophistication than distribute lists. They do not simply block or include networks like a distribute list will when a match is found, but are capable of setting the metrics on the matching route.
- **To control and modify routing information** —Route maps are used to modify routing information by setting the metrics on the matching route.
- **To define policies in policy-based routing** —Route maps make decisions based on the destination address. Once a match is found in an access list, the action is that of inclusion or exclusion. Policy-based routing allows clear decisions to be implemented on more sophisticated criteria.
- **To add granularity in the configuration of Network Address Translation (NAT)** —Route maps define pools of public and private addressing in address translation. There are additional **show** commands available by which to monitor and manage the NAT implementation.
- **To implement BGP policy-based routing** —One of the main strengths of the routing protocol BGP is its ability to perform policy-based routing. Inherent in the protocol are attributes used to affect the path taken by traffic. These are often implemented using route maps: *if* this match is made, *then* apply this attribute. This is achieved by using the **set** command to change the attributes or metric of the BGP path. In very large networks, it is important to be able to determine traffic paths. This is because of both resource and security restraints. Route maps are the main method used by BGP to define BGP routing policy.

Route maps are very similar to access lists. They both perform if/then programming, in that they state criteria that is used to determine whether specific packets are to be permitted or denied. The main difference is that the route map has the additional capability of adding a **set** action to the **match** criterion. In an access list, the match criterion is implicit; in a route map, it is a keyword. This means that if a packet is matched to the criterion given in the route map, some action can be taken to change the packet, whereas access lists can simply permit or deny the matched packet.

Until recently, you could configure a router to route traffic and place some checks and controls on the router processes or interfaces to control overhead on both the router and the network. Now, it is possible to control the nature of traffic traversing your networks. The industry has not quite achieved the full benefits of traffic engineering, but route maps provide a means by which your networks can be managed with sophistication, allowing for stable, flexible networks to grow in both size and complexity.

The characteristics of route maps are summarized in the following list:

- A route map has a list of criteria, stated with the **match** statement.
- A route map can change packets or routes that are matched by using the **set** statement.
- A collection of route map statements that have the same route map name are considered one route map.
- The route map will stop as soon as a match is made, just like an access list does.
- Within a route map, each route map statement is numbered with sequence numbers and, therefore, can be edited individually.
- The sequence number is used to specify the order in which conditions are checked. Thus, if there are two statements in a route map named BESTEST, one with sequence 5 and the other with sequence 15, sequence 5 is checked first. If there is no match for the conditions in sequence 5, then sequence 15 will be checked.
- Route maps can use IP standard or extended access lists to establish policy-based routing.
 - A standard IP access list can be used to specify match criteria for the source address of a packet.
 - Extended access lists can be used to specify match criteria based on source and destination addresses, application, protocol type, TOS, and precedence.
- The **match** route map configuration commands are used to define the conditions to be checked.
- The **set** route map configuration commands are used to define the actions to be followed if there is a match.
- A route map can contain logical AND and logical OR Boolean operations.

Like an access list, there is an implicit **deny any** at the end of a route map. The consequences of this **deny** depend on how the route map is being used.

To understand this properly, you need to see exactly how route maps operate. The following list explains the process, or logic, by which route maps work:

- The route map statements used for policy-based routing can be marked as **permit** or **deny**.
- Only if the statement is marked as **permit** and the packet meets the match criteria will the **set** commands be applied.

- The statements in a route map correspond to the lines of an access list. Specifying the match conditions in a route map is similar to specifying the source and destination addresses and masks in an access list.
- The statements in the route map are compared to the route or packet to see if there is a match. The statements are examined in turn from the top, as in an access list.
- The single **match** statement can contain multiple conditions. At least one condition in the **match** statement must be true. This is a logical OR.
- A route map statement can contain multiple **match** statements. All **match** statements in the route map statement must be considered true for the route map statement to be considered matched. This is a logical AND.

Obviously, a simple network is easier to manage and troubleshoot. Using route maps adds complexity to network management and should be handled with caution. You will learn how to configure route maps in the section “Configuring Route Maps for Policy-Based Routing,” later in this chapter.

Understanding Policy-Based Routing

Route maps are used in the configuration of policy-based routing, allowing the selection of criteria such as IP address, application, protocol, or size of packet. Once selected, the policy-based routing commands implement the policy on the selected routes.

Policy-based routes and static routes have a lot in common. However, static routes forward packets based on the destination network address, whereas a policy route forwards packets based on the source address. If access lists are used with the route map, the parameters in an extended access list can be used to route traffic based on such criteria as the destination address, length, IP protocol field, precedence, or port numbers. This gives a greater granularity and scope to the criteria by which the next-hop router is decided.

The rules that define policy-based routing are as follows:

- Traffic can be directed on either the source address or both the source and destination addresses.
- Policy-based routing affects only the routing of the router on which it is configured in determining the next hop in the path to the destination.
- Policy-based routing does not affect the destination of the packet, but it can affect the path that is taken, by setting the next hop, for example.
- Policy-based routing does not allow traffic sent into another autonomous system to take a different path from the one that would have been chosen by that autonomous system.
- It is possible to influence only how traffic will get to a neighboring router.
- As policy-based routing examines the source address, it is configured on the inbound interface.

- If there is no match made, the packet is denied policy-based routing and routed normally by destination.
- The use of route maps for policy-based routing is a little different than other applications of route maps. When used for policy-based routing, if a packet does not match the criteria specified in the route map or a matched route map statement specifies **deny**, then the packet is not dropped. It is sent to the routing process and routed normally, by destination, as if it had never encountered a route map. If your intention is to drop packets that do not match the criteria, it is necessary to use the **set** command to route packets to the null interface as the last entry in the route map.

Route maps were introduced in Cisco IOS Software Release 11.0, allowing policies that defined different paths for different packets based on specified criteria.

Policy-based routing also provides a mechanism to mark packets with different types of service (ToS). This feature can be used in conjunction with Cisco IOS queuing techniques so that certain kinds of traffic receive preferential service.

Instead of routing by the destination address, policy-based routing allows you to determine and implement routing policies to allow or deny paths based on the following:

- The identity of a particular end system
- The application being run
- The protocol in use
- The size of packets

The ability to program the path your network traffic takes adds sophistication to the routing process and the network as a whole. However, it is important to understand the benefits and disadvantages of policy-based routing, as discussed in the next sections.

Benefits of Policy-Based Routing

The benefits of implementing policy-based routing in networks include the following:

- **Source-based transit provider selection** — ISPs in particular use policy-based routing to make routing decisions based on the source address. This allows traffic belonging to different customers to be routed through different Internet connections, across the policy routers in accordance with whatever company policy needs to be adhered to.
- **Quality of service (QoS)** — By setting the precedence or type of service (TOS) values in the IP packet headers in routers at the edge of the network, organizations can provide QoS. In this way, the traffic can be differentiated, and queuing mechanisms can be implemented to prioritize

traffic based on the QoS in the core or backbone of the network. This improves network performance because the configuration is done only at the edge of the network.

- **Cost savings** — The bulk traffic generated by a specific activity can be diverted to use a higher-bandwidth, high-cost link for a short time. Meanwhile, interactive traffic is provided basic connectivity over a lower-bandwidth, low-cost link. For example, a dial-on-demand ISDN line might be raised in response to traffic to a finance server for file transfers selected by policy-based routing.
- **Load balancing** — This allows the implementation of policies to distribute traffic among multiple paths based on the traffic characteristics. This does not detract from the dynamic load-sharing capabilities offered by destination-based routing that the Cisco IOS software has always supported.

Disadvantages of Policy-Based Routing

Consider the following disadvantages before deciding to implement policy-based routing:

- A backup path should be in place in case the defined next-hop router goes down. If there is no alternative defined, policy-based routing uses the IP routing table.
- Additional CPU is required to examine every source address to effect the defined policy.
- Extra configuration is required.
- The possibility exists that other traffic will be disrupted.

Now that you understand the features of route maps and policy-based routing, the next section explains how these technologies operate together.

The Operation of Route Maps and Policy-Based Routing

As explained in the section “Understanding Route Maps,” access lists work on a simple permit and deny basis, whereas route maps can alter the characteristics of the packet or its path. For example, an access list could state something similar to this logic: If the cupcake is lemon flavored, keep it, but if it is not lemon flavored, throw it away.

Along the same lines, a route map could specify logic such as this: If it is a lemon-flavored cupcake, ice it with lemon butter frosting. If it has walnuts, then ice it with melted chocolate. If it has neither a lemon flavor nor walnuts, leave it alone. The route map is obviously more powerful than the access list because it can change the entity.

Now to show the additional complexity of route maps, add a logical AND and a logical OR. For example, if the cupcake is lemon-flavored AND it contains poppy seeds, ice it with lemon butter frosting. If it has walnuts OR it was baked today, then ice it with melted chocolate. If it does not have a lemon flavor, poppy seeds, or walnuts, leave it alone.

The route map would look something like Example 18-1.

Example 18-1 *Route Map Logic*

```
route-map cupcakes permit 10
match lemon flavored
match poppy seed
set add lemon butter frosting
route-map cupcakes permit 15
match walnuts baked today
set melted chocolate frosting
set melted chocolate frosting
route-map cupcakes permit 20
```

For the mathematicians among you, this could be written as follows:

```
If {(a and b) match} then set c
Else
If {(x or y) match} then set z
Else
Set nothing
```

Route maps are used by policy-based routing to select the packets that policy-based routing wishes to effect.

Policy-based routing is applied to *incoming* packets or packets generated by the router, if configured to do so. When a packet is received on an interface with policy-based routing enabled, it goes through this procedure:

- If there is a match and the action is to permit the route, then the packet is policy-routed in accordance to the **set** command.
- If there is a match and the action is to deny the packet, then the packet is not policy-routed but is passed back to the forwarding engine for dynamic routing.
- If there is no match and there is no configuration for what to do in this event, the default is to deny the packet, which would return it to the routing process for normal routing.
- To block packets that find no match, you need to prevent them from being returned to normal forwarding. Normal routing is prevented by specifying a **set** statement to route the packets to interface null 0 as the last entry in the route map. This will route the packets to nowhere, effectively dropping them.

Configuring Route Maps for Policy-Based Routing

This section deals with the implementation and configuration of route maps and policy-based routing. Make sure to check the Cisco documentation set for your software version before configuring a live network.

The **route-map** command is shown here:

```
Router(config)#route-map map-tag [{permit | deny} sequence-number]
```

Table 18-2 describes the syntax options available for the **route-map** command.

Table 18-2 *The route-map Command Options*

Command	Description
<i>map-tag</i>	This is the name of the route map. This name is used to reference the route map when using the ip policy route-map interface configuration command.
permit deny	(Optional) If the match criteria are met for this route map and permit is specified, the packet is forwarded as defined by the set actions. If the match criteria are not met and permit is specified, the next route map with the same map tag is tested. If there are no match criteria specified, but the packets are permitted, then all packets are set as specified. If there is no set statement, but the packets are permitted, then all packets that match the criteria are permitted. If a packet passes none of the match criteria for the set of route maps sharing the same name, it is sent to the normal routing process to be routed by destination. (Optional) If the match criteria are met for the route map and deny is specified, the packet is sent to the normal routing process, and no further route maps sharing the same map tag name will be examined.
<i>sequence-number</i>	(Optional) The sequence number indicates the position that a new route map will have in the list of route map statements already configured with the same name.

The following commands are summarized here into groups: the **match** commands that can be configured for policy-based routing, and the **set** commands that can be applied if the packet matches the criteria stated.

The match Commands for Policy-Based Routing with Route Maps

The **match** commands used in policy-based routing are summarized in Table 18-3. These **match** commands are used to determine whether the packet is to be policy-routed, as opposed to being forwarded simply by destination. If it is to be policy-routed, the packet is sent down a different path, typically one less traveled.

Table 18-3 The match Commands Used in Policy-Based Routing

Command	Description
match ip address [access-list-number name] [..access-list-number name]	This states the number or name of a standard or extended access list that will be used to examine incoming packets. A standard IP access list is used to match criteria for the source address of the packet. An extended IP access list is used to specify criteria based on source and destination, application, protocol type, TOS, and precedence. If multiple access lists are specified, matching any one will result in a match.
match length min max	This command is used to define the criteria based on the Layer 3 length of the packet. The <i>min</i> parameter states the minimum inclusive length of the packet allowed for a match. The <i>max</i> parameter states the maximum inclusive length of the packet allowed for a match. In this way, interactive traffic that is time-sensitive, such as SNA traffic tunneled in IP, can be sent on a dedicated route. Interactive traffic uses small packets, so the links could be dedicated by packet size, allowing file transfers using large packets to use a separate link so that the terminal sessions are not starved of resources.

The set Commands for Policy-Based Routing with Route Maps

The **set** commands used in policy-based routing are summarized in Table 18-4. These **set** commands are used after the **match** criteria has been satisfied. Whereas the **match** parameter determines whether the packet will be policy-routed, the **set** parameter determines how the packet is to be policy-routed.

Table 18-4 The set Commands Used in Policy-Based Routing

Command	Description
set default interface type number [...type number]	If the routing table has no explicit route for the destination network of the packet, this set provides a list of default outbound interfaces. The packet being considered for policy-based routing is routed to the available outbound interface in the list of specified default interfaces.

Table 18-4 *The set Commands Used in Policy-Based Routing (Continued)*

Command	Description
set interface <i>type number</i> [... <i>type number</i>]	If there is a route for the destination network of the packet in the routing table, this set provides a list of outgoing interfaces through which to route the packets. If more than one interface is specified, then the first functional outgoing interface is used. This command has no effect and is ignored if the packet is a broadcast or is destined to an unknown address. This is because no explicit route for the destination of the packet is found in the routing table.
set ip default next-hop <i>ip-address</i> [... <i>ip-address</i>]	If the routing table has no explicit route for the destination network of the packet, this set provides a list of default next-hop routers. The packet being considered for policy-based routing is routed to the available next hop in the list. This must be the address of an adjacent router.
set ip next hop <i>ip-address</i> [... <i>ip-address</i>]	This set provides a list of next-hop routers to which to forward the packet. If more than one next hop is specified, then the first available next-hop router is used. This must be the address of an adjacent router, and there must be an entry for the destination network of the packet in the routing table.
set ip precedence <i>precedence</i>	This is used to set the precedence bits in the Type of Service field of the IP header of the matched packet. This determines the IP precedence in the IP packets.
set ip tos <i>type-of-service</i>	This is used to set the IP ToS value in the Type of Service field of the IP header.

The **set** commands can be used in conjunction with each other.

Once configured, the route map must be called into service. Until it is called, it has no power. The command used to recruit the services of the route map to an incoming interface follows:

```
Router(config-if)#ip policy route-map map-tag
```

map-tag is the name of the route map to use for policy-based routing. This must match a map tag specified by a **route-map** command.

Policy-based routing is configured on the incoming interface that receives the packets and performs policy-based routing on incoming packets, determining the path of the packet to the destination.

With the appropriate configuration, you can apply policy-based routing on packets generated by the router. The command is configured globally, using the following syntax:

```
Router(config)# ip local policy route-map
```

Example 18-2 shows a sample configuration.

Example 18-2 *Calling a Route Map into Service*

```
Router(config)#interface serial 0
Router(config-if)#ip policy route-map soup spoon
!
Router(config)#route-map soup spoon permit 10
Router(config-route-map)#match ip address 1
Router(config-route-map)#set ip next-hop 191.5.6.11
Router(config)#access-list 1 permit 201.14.222.18
```

There are many things to be aware of when configuring a router that is directing the network traffic. When configuring policy-based routing or route maps, pay very careful attention to the logic and rules by which they operate.

CAUTION When editing a route map statement with the **no** version of the existing command line, if you forget to type in the sequence number, you will delete the entire route map.

Configuring Fast Switching with Policy-Based Routing

Speed through the network is influenced by the capability of the network devices to process traffic. Cisco is continually striving to enhance the features of its products, while at the same time reducing the resources consumed and the time it takes to provide those features.

Cisco made a major achievement in Cisco IOS Software Release 11.2F. In this version of IOS, IP policy-based routing is fast-switched. The previous versions' process-switch policy-routed traffic, allowing for an output of approximately 1000 to 10,000 packets per second, resulted in application timeouts.

Fast switching of policy-based routing is disabled by default. You must configure it manually. To do so, complete the following steps:

- Step 1** Configure policy-based routing before you configure fast-switched policy-based routing.
- Step 2** When policy-based routing is configured, turn on the fast switching with this interface command:

```
Router(config-if)#ip route-cache policy
```

Fast-switched policy-based routing supports all of the **match** commands and most of the **set** commands, except for the following restrictions:

- The **set ip default** command is not supported.
- The **set interface** command is supported only over point-to-point links, unless a route-cache entry exists using the same interface specified in the **set interface** command in the route map. The route cache is the portion of memory assigned to the product of routing decisions. In addition, when process switching, the routing table is consulted to determine a path to the destination. During fast switching, the software does not make this check because fast switching is a cache of the process switch lookup. Instead, if the packet matches, the software blindly forwards the packet to the specified interface. This is a similar situation to the one described in reference to load balancing earlier in the section “Benefits of Policy-Based Routing.”

This next section expands the discussion of route maps to include redistribution.

Configuring Route Maps for Redistribution

Although the filtering discussed in Chapter 17, “Implementing Redistribution and Controlling Routing Updates,” is perfectly adequate for simply denying or permitting routes from entering another routing process, route maps can do more. Their strength lies in their ability to change the route in some way. A common manipulation of the route using route maps is to change the metric. As you saw in Chapter 17, changing the metric is necessary so that the receiving routing protocol can forward the route using a metric that it understands.

The following commands are summarized here into groups: the **match** commands that can be configured for redistribution, and the **set** commands that can be applied if the route matches the criteria stated.

The match Commands for Redistribution with Route Maps

The **match** commands used in redistribution are summarized in Table 18-5. These **match** commands are used to determine whether the route is to be redistributed.

Table 18-5 *The match Commands Used in Redistribution*

Command	Description
match interface (IP)	Distributes any routes that have their next hop out one of the interfaces specified.
match ip address [<i>access-list-number</i> <i>name</i>] [... <i>access-list-number</i> <i>name</i>]	This is the same command as used in policy-based route maps and serves the same function as described in Table 18-3. The command states the number or name of a standard or extended access list that will be used to examine incoming packets.

(continues)

Table 18-5 The **match** Commands Used in Redistribution (Continued)

Command	Description
match ip next-hop	Redistributes any routes that have a next hop router address passed by one of the access lists specified.
match ip route-source	Redistributes routes that have been advertised by routers and access servers at the address specified by the access lists.
match metric	Redistributes routes with the metric specified.
match route-type (IP)	Redistributes routes of the specified type.
match tag	Redistributes routes in the routing table that match the specified tags.

NOTE If a route is not matched, it is not redistributed.

The **set** Commands for Redistributing with Route Maps

The following **set** commands are used after the **match** criteria have been satisfied. Whereas the **match** parameter determines whether the route will be redistributed, the **set** parameter determines how the route is to be redistributed.

The **set** command is as follows:

```
Router(config-route-map)#set {criteria}
```

The **set** commands used in redistribution are summarized in Table 18-6.

Table 18-6 The **set** Commands Used in Redistribution

Command	Description
set level { <i>level-1</i> <i>level-2</i> <i>level-1-2</i> <i>stub-area</i> <i>backbone</i> }	Used by IS-IS to determine the level of router to which the process should import routes. Also used by OSPF to state the type of area router to which routes should be imported.
set metric (BGP, OSPF, RIP)	Sets the metric value for a routing protocol.
set metric-type { <i>internal</i> <i>external</i> <i>type-1</i> <i>type-2</i> }	Sets the metric type for the destination routing protocol.
set tag <i>tag-value</i>	Sets a tag value of the destination routing protocol.

Once configured, the route map must be called into service. Until it is called, it has no power. The command used to recruit the services of the route map for redistribution is the redistribution command itself. Once configured, redistribution sends routes to the route map.

```
Router(config-router)#redistribute protocol [process-id] [route-map map-tag]
```

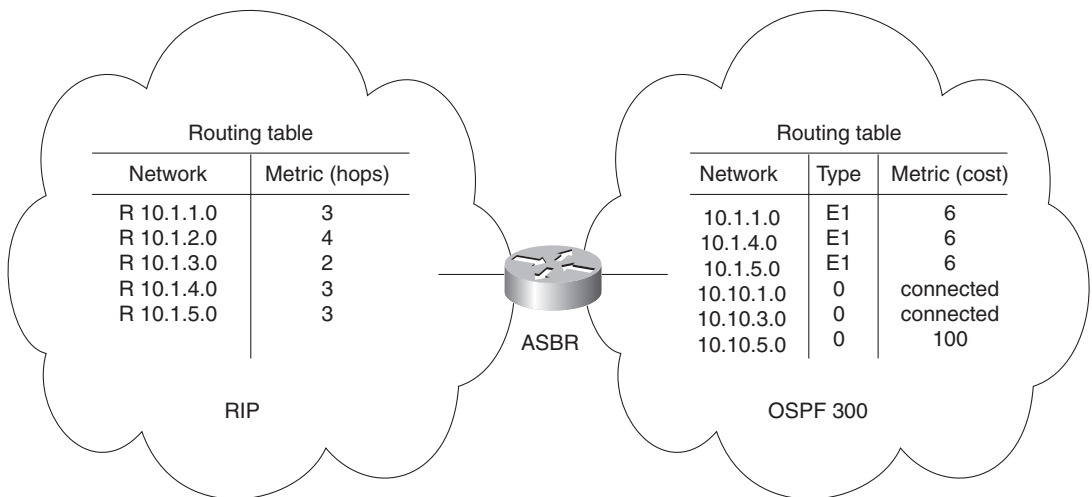
map-tag is the name of the route map to use for redistribution. This must match a map tag specified by a **route-map** command.

Example 18-3 is very simple, but it clearly illustrates the functionality of the route map. Study the example in reference to Figure 18-1.

This route map examines all updates from RIP and redistributes those RIP routes with a hop count equal to 3 into OSPF. These routes will be redistributed into OSPF as external link-state advertisements (LSAs) with a metric cost of 6, a metric type of Type 1, and a tag equal to 1.

The tag is useful for tracking routes during redistribution, when the routes change from one routing domain to another, for example, from RIPv2 to OSPF. The routes are tagged at the point at which they are redistributed into another protocol. Although the routing protocols do not use the tags, they are passed between the different domains during redistribution.

Figure 18-1 Route Map to Distribute RIPv2 into OSPF



Example 18-3 Route Map to Distribute RIPv2 into OSPF

```
Router(config)#router ospf 25
Router(config-router)#redistribute rip route-map rip-routes

Router(config)#route-map rip-routes permit 10
Router(config-route-map)#match metric 3
Router(config-route-map)#set metric 6
Router(config-route-map)#set metric-type type-1
Router(config-route-map)#set tag 1
```

Monitoring the Configuration of Route Maps, Policy-Based Routing, and Redistribution

Most of the appropriate commands in tracking route maps are the same as those shown in Chapter 17. The commands used to test connectivity throughout the network include the following:

- **show ip protocol**
- **show ip route**
- **show ip route** *routing-protocol*
- **show ip eigrp neighbors**
- **show ip ospf database**

In addition to these commands, **trace** and extended **ping** are also very useful. **ping** is particularly useful in policy-based routing where packets are routed based on packet length.

To monitor the policy-based-routing configuration, use the following EXEC commands described in Table 18-7.

Table 18-7 *Commands to Monitor Policy-Based Routing*

Command	Description
show ip policy	Displays the route maps used for policy-based routing on the router’s interfaces.
show route-map	Displays configured route maps.
debug ip policy	Displays IP policy-based-routing packet activity. This command helps you to determine what policy-based routing is doing. It displays information about whether a packet matches the criteria and, if so, the resulting routing information for the packet.

CAUTION Because the **debug ip policy** command generates a significant amount of output, use it only when traffic on the IP network is low so that other activity on the system is not adversely affected. This is true of all **debug** commands.

Foundation Summary

The “Foundation Summary” section of each chapter lists the most important facts from the chapter. Although this section does not list every fact from the chapter that will be on your exam, a well-prepared candidate should, at a minimum, know all the details in each “Foundation Summary” before going to take the exam.

The characteristics of route maps are summarized in the following list:

- A route map has a list of criteria, stated with the **match** statement.
- A route map can change packets or routes that are matched by using the **set** statement.
- A collection of route map statements that have the same route map name are considered one route map.
- The route map will stop as soon as a match is made, just like an access list does.
- Within a route map, each route map statement is numbered with sequence numbers and, therefore, can be edited individually.
- Route maps can use IP standard or extended access lists to establish policy-based routing.
 - A standard IP access list can be used to specify match criteria for the source address of a packet.
 - Extended access lists can be used to specify match criteria based on source and destination addresses, application, protocol type, TOS, and precedence.
- The **match** route map configuration commands are used to define the conditions to be checked.
- The **set** route map configuration commands are used to define the actions to be followed if there is a match.
- A route map can contain logical ANDs as well as logical ORs.
- The sequence number is used to specify the order in which conditions are checked. Thus, if there are two statements in a route map named BESTEST, one with sequence 5 and the other with sequence 15, sequence 5 is checked first. If there is no match for the conditions in sequence 5, then sequence 15 will be checked.

The following characterize the operation of route map statements:

- The route map statements used for policy-based routing can be marked as **permit** or **deny**.
- Only if the statement is marked as **permit** and the packet meets the match criteria will the **set** commands be applied.

- The statements in a route map correspond to the lines of an access list. Specifying the match conditions in a route map is similar to specifying the source and destination addresses and masks in an access list.
- The statements in the route map are compared to the route or packet to see if there is a match. The statements are examined in turn from the top, like in an access list. The single **match** statement can contain multiple conditions. At least one condition in the **match** statement must be true. This is a logical OR.
- A route map statement can contain multiple **match** statements. All **match** statements in the route map statement must be considered true for the route map statement to be considered matched. This is a logical AND.

The **route-map** command syntax is shown here:

```
Router(config)#route-map map-tag [{permit | deny} sequence-number]
```

Table 18-8 describes the syntax options available for the **route-map** command.

Table 18-8 *The route-map Command Options*

Command	Description
<i>map-tag</i>	This is the name of the route map. This name is used to reference the route map when applied as an interface configuration command for policy-based routing or when called by the routing process in redistribution.
permit deny	(Optional) If the match criteria are met for this route map and permit is specified, the packet is forwarded as defined by the set actions. If the match criteria are not met and permit is specified, the next route map with the same map tag is tested. If there are no match criteria specified, but the packets or routes are permitted, then all packets or routes are set as specified. If there is no set statement, but the packets or routes are permitted, then all packets or routes that match the criteria are permitted. If a packet passes none of the match criteria for the set of route maps sharing the same name, it is sent to the normal routing process to be routed by destination. (Optional) If the match criteria are met for the route map and deny is specified, the packet is sent to the normal routing process, and no further route maps sharing the same map tag name will be examined.
<i>sequence-number</i>	(Optional) The sequence number indicates the position that a new route map will have in the list of route map statements already configured with the same name.

The **match** commands used in policy-based routing and redistribution are summarized in Table 18-9.

Table 18-9 *The match Commands Used in Policy-Based Routing*

Command	Description
match interface (IP)	Distributes any routes that have their next hop out one of the interfaces specified.
match ip address [<i>access-list-number</i> <i>name</i>] [... <i>access-list-number</i> <i>name</i>]	This states the number or name of a standard or extended access list that will be used to examine incoming packets. A standard IP access list is used to match criteria for the source address of the packet. An extended IP access list is used to specify criteria based on source and destination, application, protocol type, TOS, and precedence. If multiple access lists are specified, matching any one will result in a match.
match ip next-hop	Redistributes any routes that have a next-hop router address passed by one of the access lists specified.
match ip route-source	Redistributes routes that have been advertised by routers and access servers at the address specified by the access lists.
match length <i>min max</i>	This command is used to define the criteria based on the Layer 3 length of the packet. The <i>min</i> parameter states the minimum inclusive length of the packet allowed for a match. The <i>max</i> parameter states the maximum inclusive length of the packet allowed for a match. In this way, interactive traffic that is time-sensitive, such as SNA traffic tunneled in IP, can be sent on a dedicated route. Interactive traffic uses small packets, so the links could be dedicated by packet size, allowing file transfers using large packets to use a separate link so that the terminal sessions are not starved of resources.
match metric (IP)	Redistributes routes with the metric specified.
match route-type (IP)	Redistributes routes of the specified type.
match tag	Redistributes routes in the routing table that match the specified tags.

The **set** commands used in policy-based routing are summarized in Table 18-10.

Table 18-10 *The set Commands Used in Route Maps*

Command	Description
set automatic-tag	Automatically computes the tag value.
set default interface <i>type number</i> [... <i>type number</i>]	If the routing table has no explicit route for the destination network of the packet, this set provides a list of default outbound interfaces. The packet being considered for policy-based routing is routed to the available outbound interface in the list of specified default interfaces.
set interface <i>type number</i> [... <i>type number</i>]	If there is a route for the destination network of the packet in the routing table, this set provides a list of outgoing interfaces through which to route the packets. If more than one interface is specified, then the first functional outgoing interface is used. This command has no effect and is ignored if the packet is a broadcast or is destined to an unknown address. This is because no explicit route for the destination of the packet is found in the routing table.
set ip default next-hop <i>ip-address</i> [... <i>ip-address</i>]	If the routing table has no explicit route for the destination network of the packet, this set provides a list of default next-hop routers. The packet being considered for policy-based routing is routed to the available next hop in the list. This must be the address of an adjacent router.
set ip next hop <i>ip-address</i> [... <i>ip-address</i>]	If there is a route for the destination network of the packet in the routing table, this set provides a list of next-hop routers to which to forward the packet. If more than one next hop is specified, then the first available next-hop router is used. This must be the address of an adjacent router.
set ip precedence <i>precedence</i>	This is used to set the precedence bits in the Type of Service field of the IP header of the matched packet. This determines the IP precedence in the IP packets.
set ip tos <i>type-of-service</i>	This is used to set the IP ToS value in the Type of Service field of the IP header.
set level { <i>level-1</i> <i>level-2</i> <i>level-1-2</i> <i>stub-area</i> <i>backbone</i> }	Used by IS-IS to determine the level of router to which the process should import routes. Also used by OSPF to state the type of area router to which routes should be imported.
set metric (BGP, OSPF, RIP)	Sets the metric value for a routing protocol.
set metric-type { <i>internal</i> <i>external</i> <i>type-1</i> <i>type-2</i> }	Sets the metric type for the destination routing protocol.
set tag <i>tag-value</i>	Sets a tag value of the destination routing protocol.

Figure 18-2 and Figure 18-3 summarize the logic used when route maps are applied to a router. Remember that in Figure 18-3, if there is no match or the match is a deny, the packet is not discarded but sent to the routing process to be routed by destination. If, however, the routing table has no entry for the destination, the packet will be dropped. This is not a function of route maps but rather the normal routing process.

Figure 18-2 *Route Map Logic for Policy-Based Routing 1*

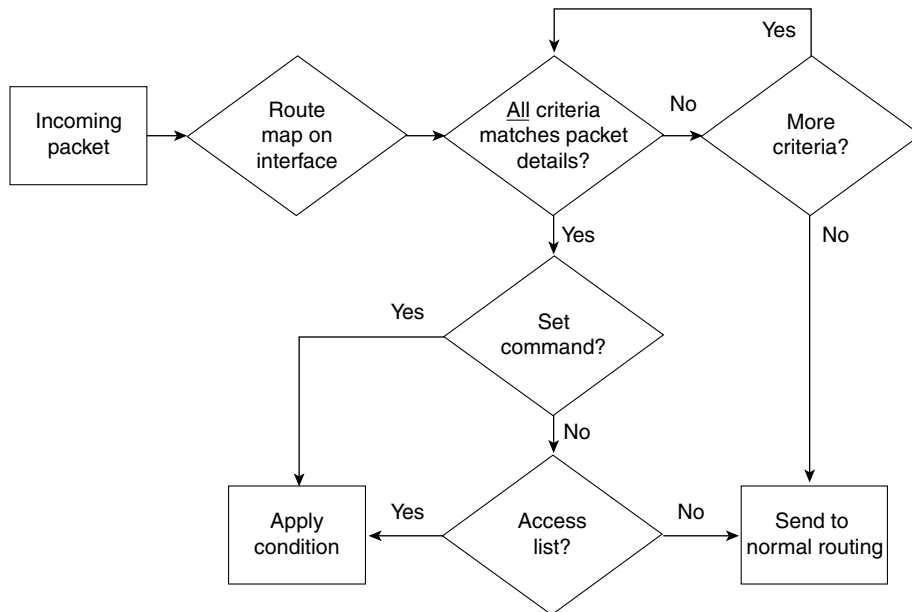
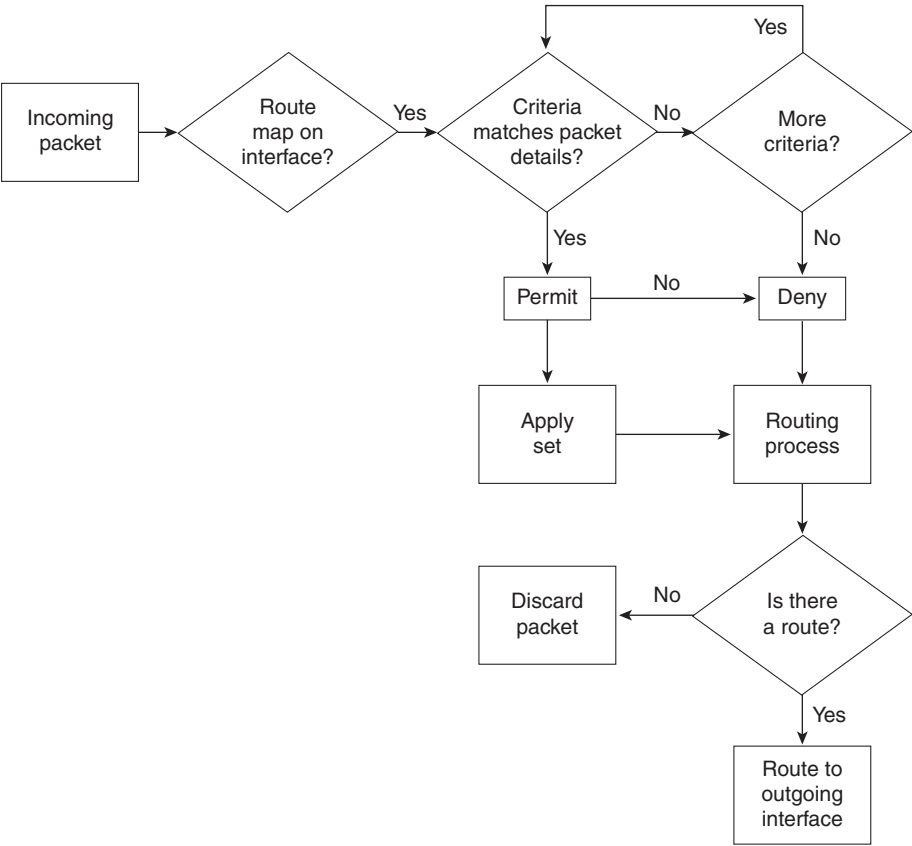


Figure 18-3 *Route Map Logic for Policy-Based Routing 2*



Q&A

As mentioned in the introduction, “All About the CCNP, CCDP, and CCIP Certifications,” you have two choices for review questions. The questions that follow next give you a bigger challenge than the exam itself by using an open-ended question format. By reviewing now with this more difficult question format, you can exercise your memory better and prove your conceptual and factual knowledge of this chapter. The answers to these questions are found in Appendix A.

For more practice with examlike question formats, including questions using a router simulator and multichoice questions, use the exam engine on the CD-ROM.

1. Explain the command **match** *ip address {access-list-number | name} [access-list number | name]*.
2. Explain the command **ip route-cache policy** .
3. State two benefits of using policy-based routing.
4. How are matching routes modified in a route map?
5. Explain the command **set ip default next-hop** [*ip-address...ip-address*].
6. Which command displays route maps that are configured on interfaces?
7. What command is used to attach a route map to an incoming interface?
8. What is a map tag?
9. Explain briefly the difference between the **match** and **set** commands.
10. What are the criteria by which policy-based routes are determined?
11. How would you block traffic that found no match?
12. What is the purpose of the sequence number in a route map?
13. What logic is used if there are multiple match statements in the route map?
14. What parameters can an extended access list define in the route map selection process?
15. How can policy-based routing save money for the network?
16. What are some of the potential disadvantages of using policy-based routing?
17. Explain the use of the command **match length** *min max*.
18. Explain the following sample configuration.
19. Briefly explain the use of the command **show ip policy** .
20. Configuring route maps is complex, and it is easy to confuse the logic by which they work. State one of the things you should be aware of when configuring a route map.

Scenarios

The following scenarios and questions are designed to draw together the content of the chapter and to exercise your understanding of the concepts. There is not necessarily a right answer. The thought process and practice in manipulating the concepts are the goals of this section. The answers to the scenario questions are found at the end of this chapter.

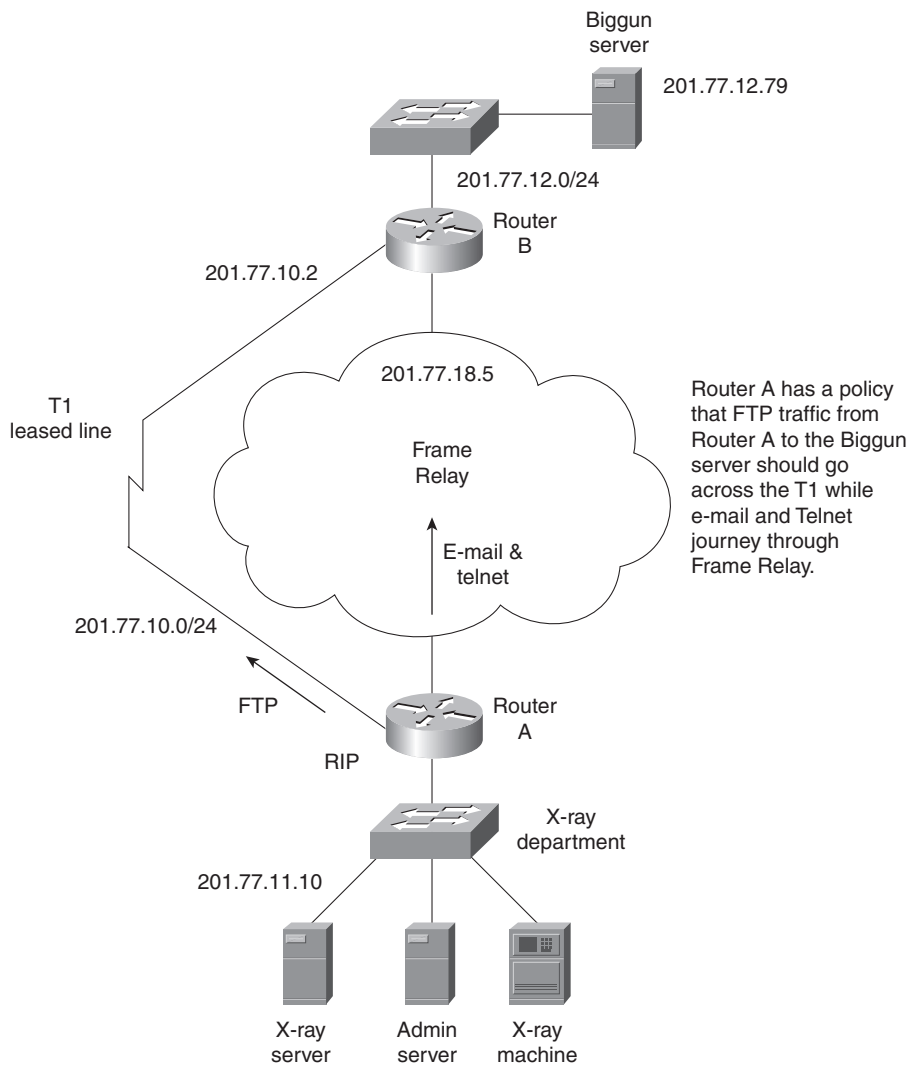
Scenario 18-1

The hospital Duddleduddle needs to implement a routing policy using route maps. This is to ensure the optimum use of bandwidth. The X-ray department requires an enormous amount of bandwidth when transferring MRI (Magnetic Resonance Imaging) images to the centralized database.

Using Figure 18-4 and given the criteria within the questions, configure the route maps on Router A.

1. The hospital policy states that the FTP traffic from the X-ray department (201.77.11.0/24) should be forwarded to the Biggun Server at 201.77.12.79 and that it should be sent across the leased line, which is a T1 connection. What would the configuration look like?
2. The Telnet sessions and e-mail connections should be sent across the Frame Relay link. This traffic is from the same department (201.77.11.0.0/24) and is connecting to the same server. What would the configuration look like?
3. What commands would you use to verify that the policy-based routing is configured correctly and operating normally?

Figure 18-4 *Route Maps Example for Scenario 18-1*



Scenario Answers

The answers provided in this section are not necessarily the only possible answers to the questions. The questions are designed to test your knowledge and to give practical exercise in certain key areas. This section is intended to test and to exercise skills and concepts detailed in the body of this chapter.

If your answer is different, ask yourself whether it follows the tenets explained in the answers provided. Your answer is correct not if it matches the solution provided in the book, but rather if it has included the principles of design laid out in the chapter.

In this way, the testing provided in these scenarios is deeper: It examines not only your knowledge, but also your understanding and ability to apply that knowledge to problems.

If you do not get the correct answer, refer back to the text and review the subject tested. Be certain to also review your notes on the question to ensure that you understand the principles of the subject.

Scenario 18-1 Answers

1. *The hospital policy states that the FTP traffic from the X-ray department (201.77.11.0/24) should be forwarded to the Biggun Server at 201.77.12.79 and that it should be sent across the leased line, which is a T1 connection. What would the configuration look like?*

The configuration would be as follows:

```
Router(config)#interface e0
Router(config-if)#ip address 201.77.11.1 255.255.255.0
Router(config-if)#ip policy route-map xray

Router(config)#access-list 101 permit tcp 201.77.11.0 0.0.0.255 any eq ftp
Router(config)#access-list 101 permit tcp 201.77.11.0 0.0.0.255 eq ftp-data
any
Router(config)#route-map xray permit 10
Router(config)#match ip address 101
Router(config)#set ip next-hop 201.77.10.2
```

2. *The Telnet sessions and e-mail connections should be sent across the Frame Relay link. This traffic is from the same department (201.77.11.0/24) and is connecting to the same server. What would the configuration look like?*

The configuration would be as follows:

```
Router(config)#interface e0
Router(config-if)#ip address 201.77.11.1 255.255.255.0
Router(config-if)#ip policy route-map xray

Router(config)#access-list 101 permit tcp 201.77.11.0 0.0.0.255 any eq ftp
Router(config)#access-list 101 permit tcp 201.77.11.0 0.0.0.255 eq ftp-data
any
Router(config)#access-list 106 permit tcp 201.77.11.0 0.0.0.255 any eq smtp
Router(config)#access-list 106 permit tcp 201.77.11.0 0.0.0.255 any eq telnet

Router(config)#route-map xray permit 10
Router(config-route-map)#match ip address 101
Router(config-route-map)#set ip next-hop 201.77.10.2
Router(config)#route-map xray permit 20
Router(config-route-map)#match ip address 106
Router(config-route-map)#set ip next-hop 201.77.18.5
```

Remember that if there is no match or the match is a deny, the packet is not discarded but sent to the routing process to be routed by destination. If, however, the routing table has no entry for the destination, the packet will at this point be dropped. This is not a function of route maps but rather of the normal routing process.

3. *What commands would you use to verify that the policy-based routing is configured correctly and operating normally?*

The commands that should be used to verify the policy-based routing are as follows:

- **show ip policy**
- **show route-map** *name*
- **debug ip policy**

Part VII: Scenarios

Chapter 19 Scenarios for Exam Preparation

Part VII covers the following Cisco BSCI exam topics:

- Given specific requirements, choose the correct routing protocol to meet the requirements
- Describe the concepts relating to route summarization and apply them to hypothetical scenarios
- Given a set of network requirements, identify the steps to configure an EIGRP environment and verify proper operation (within described guidelines) of your routers
- Given an addressing scheme and other laboratory parameters, identify the steps to configure a multiple-area OSPF environment and verify proper operation (within described guidelines) of your routers
- Identify the steps to configure route redistribution in a network
- Identify the steps to select and configure the different ways to control routing update traffic
- Describe concepts relating to extending IP addresses and the use of VLSMs to extend IP addresses
- Describe the features and operation of EIGRP
- Describe the features and operation of multiarea OSPF
- Compare classful and classless routing protocols



Scenarios for Exam Preparation

This chapter is designed to assist you in final preparation for the BSCI exam by providing additional practice with the core topics of the exam. These exercises and tasks require a broad perspective, which means that you will need to draw on knowledge that you acquired in Chapters 1 through 18. This chapter also focuses on configuration and verification commands. These scenarios are designed with certain assumptions about the way we all learn and retain information.

It is easy to forget the details of a particular technology by the time you have completed your study of the other chapters. To resolve this, the scenarios in this chapter cover the entire breadth of exam topics to remind you about many of these details. In addition to simply covering the exam topics, the scenarios put the topics in context. This makes it easier to comprehend each subject because it is not presented in a vacuum. After successfully working through the scenarios, you can feel confident in your preparedness not only for the exam, but also for real-life implementation of the technologies.

The ability to apply knowledge in different contexts proves you have achieved a deep understanding of the subject matter. These scenarios challenge you outside the scope of the individual facts that are presented, which makes those facts much easier to remember and apply.

Your understanding of the concepts at this point in your study is complete; practice and repetition is useful so that you can answer quickly and confidently on the exam. Good luck.

Further Study for Final Exam Preparation

This chapter is not the only chapter that you should use when doing your final preparation for the BSCI exam. Not all the subjects in the exam are covered in this chapter. Here is a brief list of the study options provided by this book, beyond the core chapters and this chapter:

- All prechapter quiz and chapter-ending questions, with answers, are in Appendix A, “Answers to Chapter ‘Do I Know This Already?’ Quizzes and Q&A Sections.” These conveniently located questions can be read and reviewed quickly.

- The CD-ROM contains practice exam questions that you can use to take an overall sample exam or test yourself on specific topics. The CD-ROM also contains an electronic version of the book in addition to supplemental material.
- Each core chapter has a “Foundation Summary” section near the end that contains concise tables and information for final review.
- The Glossary is also a good study aid.

How to Best Use This Chapter

This chapter includes two types of scenarios:

- **Basic scenarios** — Standard questions based on a presented situation. The scenario draws together technology and implementation issues that are broader than those presented in the specific subject chapters. These are Scenarios 19-1, 19-2, and 19-3. The answers to these scenarios follow the three scenarios.
- **Three-part scenarios** — More complicated scenarios that require planning, configuration, and verification. These are Scenarios 19-4, 19-5, and 19-6. The solutions to the three-part scenarios are contained within each scenario.

The scenarios focus on easily forgotten items, for example, the **show** and **debug** commands. Their options are often ignored, mainly because you can get online help about the correct option easily when using the Cisco command-line interface (CLI). However, questions about the exact command options you use to see a particular piece of information are scattered throughout the exam. Take care to review the output of the commands in these scenarios.

Scenarios

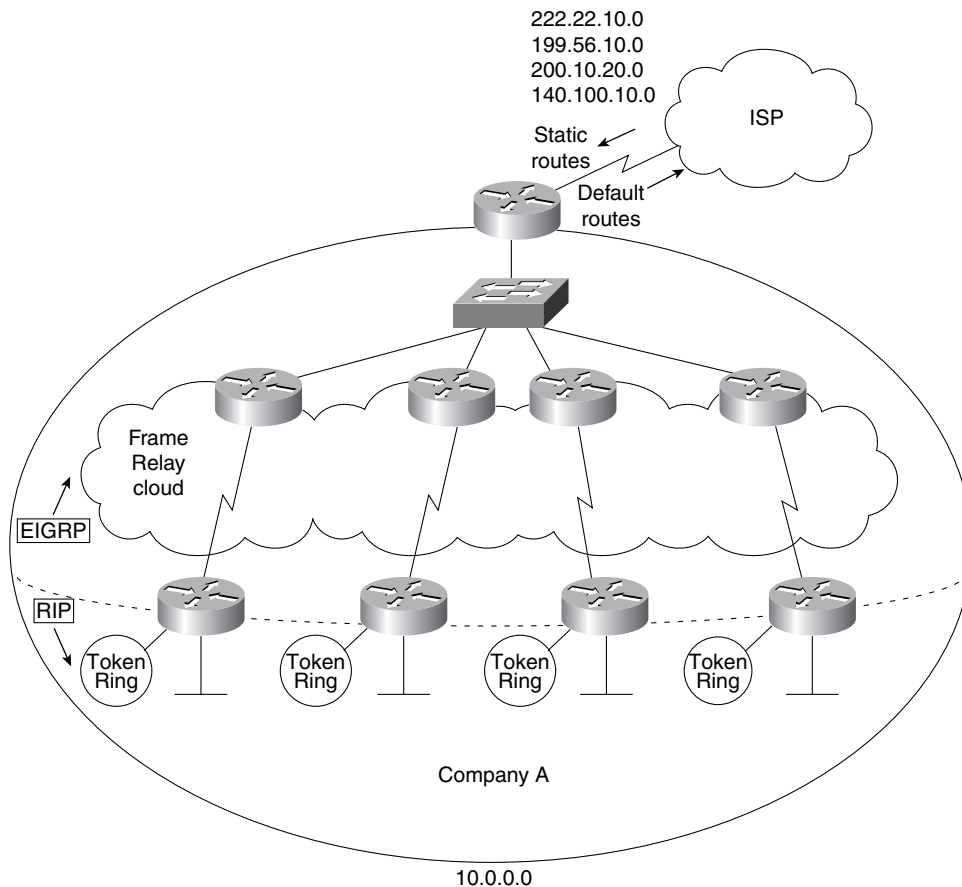
The following scenarios and questions are designed to draw together the content of the chapters and to exercise your understanding of the concepts. There is not necessarily a right answer. The thought process and practice in manipulating the concepts are the goals of this section. The answers to these scenarios immediately follow Scenario 19-3.

Scenario 19-1

The last network administrator of your company left abruptly. You were recently hired to the position and cannot find any documentation on the network. Using a network-management tool, you now have a topology map of the network.

Refer to Figure 19-1 and answer the questions that follow.

1. Offer some reasons why the routing protocol EIGRP has recently been implemented in the network, and give the reasons for its deployment in that particular area of the network.
2. The ISP has decided that the company should set a default route into its domain. Write out the commands that would configure the default route.
3. Is the router that is connecting to the ISP the only router that needs configuration? If this is the case, explain the reasons for your decision. If other routers need configuration, explain why they are needed and state where they would be applied.
4. The ISP router connecting into the company network will need static routes configured. Write out the commands to configure the router.
5. Will these routes need to be redistributed into the ISP domain? If the answer is yes, explain the reason for your decision, and then write out the configuration commands. If the answer is no, explain the reasons for your decision.
6. Are filters required in this design? If so, state why, describe the type of filters required, and explain how and where they would be applied.
7. A consultant has suggested that EIGRP should be configured using summarization. If this plan were to be implemented, on which routers would it be configured, and why?

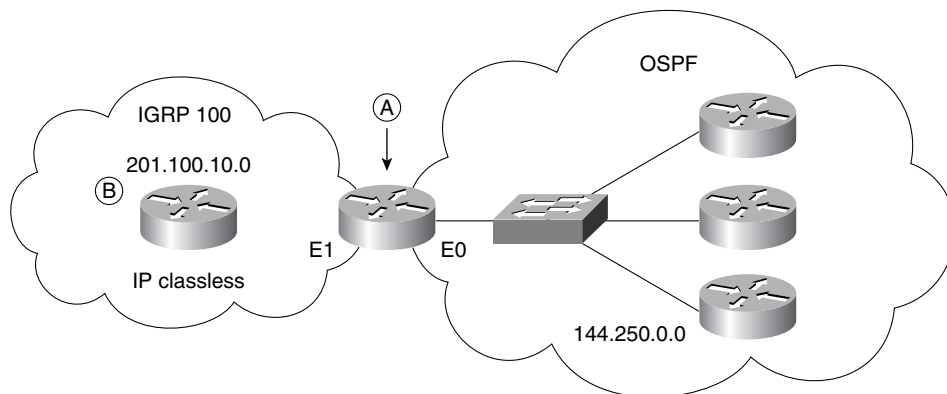
Figure 19-1 *Diagram of Network for Scenario 19-1*

Scenario 19-2

The company has recently merged with another company that has OSPF as its routing protocol. It has been determined that this is going to be the company standard. Figure 19-2 shows the topology for this scenario.

1. Router A was chosen as the designated router on the LAN in the OSPF domain. Explain why this design choice was made, and give the command that would ensure its selection as the designated router.
2. Redistribution is occurring between IGRP and OSPF on Router A; state the configuration commands that you would use for both IGRP and OSPF.

Figure 19-2 Diagram for Scenario 19-2



3. The new headquarters of the merged company is in the OSPF domain. Therefore, because the connection to the Internet is also in this domain, it has been decided that Router A will have a default route configured to point to the OSPF domain from the IGRP domain. It will be necessary to configure the **ip classless** command for the IGRP domain. Explain why the **ip classless** command is needed and what it achieves.
4. Explain the use of the **subnet** command in OSPF, and why it is required in redistribution.
5. In a more complex topology with multiple redistribution points, explain why it might be necessary to filter when redistributing, and what it achieves in this network.

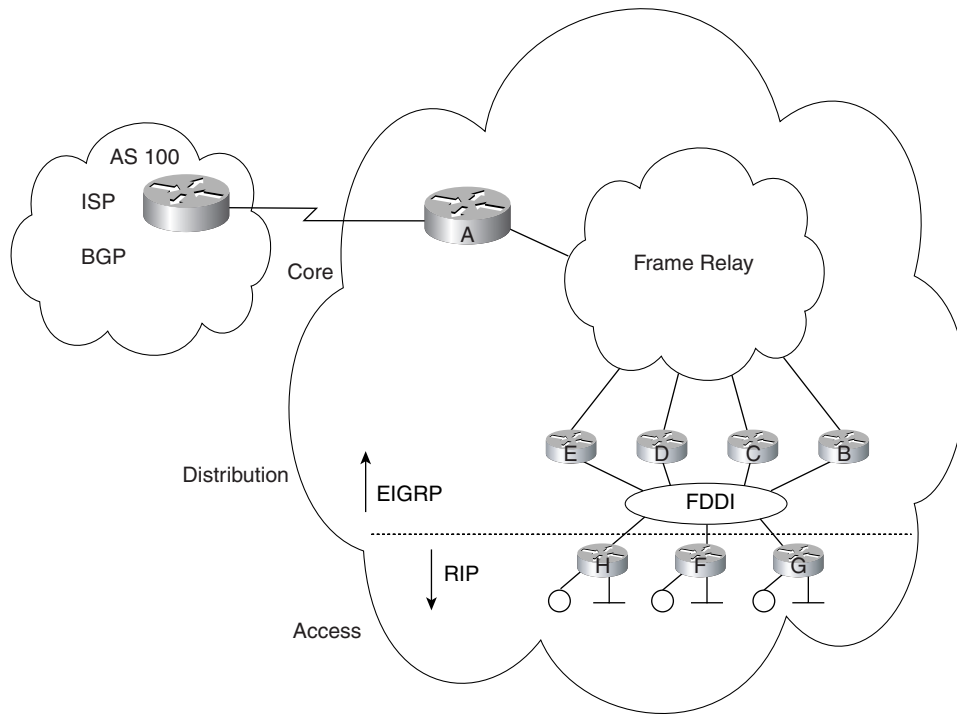
Scenario 19-3

Your company has decided to change its routing protocol from RIP to EIGRP. The company is currently running both IPX and IP. With reference to Figure 19-3, devise a design for the new network.

Ensure that you include the following:

1. Create a transition plan, showing how the final design will be implemented.
2. If the transition plan requires redistribution, indicate these points on the diagram and provide the configuration commands.
3. The company has also decided to connect to the Internet using EIGRP, upon which the ISP that is providing Internet connectivity has agreed. Is any redistribution required? Give reasons for your answers.

Figure 19-3 *Diagram of Network for Scenario 19-3*



4. Explain the neighbors that you see on the LAN segment and the path selection on Router B to network 140.100.0.0.
5. Create a diagram that shows the final network design, with configuration commands for EIGRP on Router A.
6. Write out the configuration commands for Router B.

Scenario Answers

The answers provided in this section are not necessarily the only possible answers to the questions. The questions are designed to test your knowledge and to give practical exercise.

If your answer is different, ask yourself whether it follows the tenets explained in the answers provided. Your answer is correct not if it matches the solution provided in the book, but rather if it has included the principles of design laid out in the chapter.

In this way, the testing provided in these scenarios is deeper: The scenarios examine not only your knowledge, but also your understanding and ability to apply that knowledge to problems.

If you do not get the correct answer, refer back to the previous chapters and review the subject tested. Be certain to also review your notes on the question to ensure that you understand the principles of the subject.

Scenario 19-1 Answers

1. *Offer some reasons why the routing protocol EIGRP has recently been implemented in the network, and give the reasons for its deployment in that particular area of the network.*

EIGRP is being run across the Frame Relay WAN in areas using incremental updates, which is a more efficient use of resources. The network routers are all Cisco devices capable of running the proprietary protocol EIGRP.

Using EIGRP is far more efficient on a WAN because incremental updates can be sent across the limited bandwidth. In particular, if either AppleTalk or IPX is running at the access level, the routing updates can be sent in EIGRP, which gives control and flexibility. EIGRP is used because it converges quickly, is classless, and is easy to deploy and maintain.

2. *The ISP has decided that the company should set a default route into its domain. Write out the commands that would configure the default route.*

The configuration commands to establish a default route from the company's network to the ISP domain are as follows:

```
Router(config)# router eigrp 100  
Router(config-router)# network 10.0.0.0  
Router(config-router)# network 140.100.10.0  
Router(config)# ip default-network 140.100.0.0
```

3. *Is the router that is connecting to the ISP the only router that needs configuration? If this is the case, explain the reasons for your decision. If other routers need configuration, explain why they are needed and state where they would be applied.*

All the routers require a basic EIGRP configuration, but the router connecting to the ISP is the only router that needs additional configuration to connect to the ISP. Redistribution between the EIGRP processes occurs within the ISP. EIGRP automatically creates a default route, which will be propagated to all the other EIGRP routers automatically. The routers connecting to the RIP network need the EIGRP routes redistributed. Default routes redistributed into an RIP environment might require the **ip classless** command to be configured; this is for routers that are downstream from the router that is dealing with the redistribution. Because the routers at the redistribution points will have routes redistributed into them by EIGRP, they will have no problem in the lookup.

4. *The ISP router connecting into the company network will need static routes configured. Write out the commands to configure the router.*

The configuration commands to establish static routes from the ISP domain into the company's network are as follows:

```
Router(config)# ip route 140.100.10.0 255.255.255.0 140.100.60.3
Router(config)# ip route 200.10.20.0 255.255.255.0 140.100.60.3
Router(config)# ip route 199.56.10.0 255.255.255.0 140.100.60.3
Router(config)# ip route 222.22.10.0 255.255.255.0 140.100.60.3
```

5. *Will these routes need to be redistributed into the ISP domain? If the answer is yes, explain the reason for your decision, and then write out the configuration commands. If the answer is no, explain the reasons for your decision.*

The static commands defined in the previous answer need to be redistributed into the ISP domain. Static routes are not redistributed automatically. The **passive-interface** command prevents EIGRP updates from being sent to the customer.

The commands for redistribution of static routes are as follows:

```
Router(config)# router eigrp 100
Router(config)# redistribute static
Router(config)# default-metric 10000 100 255 1 1500
Router(config)# passive-interface s0
```

6. *Are filters required in this design? If so, state why, describe the type of filters required, and explain how and where they would be applied.*

Because there are no redundant paths in the redistribution between protocols, no filters are required in the configuration of the routers. There will be no feedback between the protocols.

Because there is no routing protocol running between the company and the ISP, there is no requirement for filters here, either. If any filters were required, they would be for internal security and traffic control and would be typically configured on the access routers entering the

Frame Relay cloud. There could also be some distribute lists at the distribution layer to limit connectivity among the different regions.

7. *It has been suggested that EIGRP should be configured using summarization. If this plan were to be implemented, on which routers would it be configured, and why?*

Summarization would be configured at the access level to limit the number of updates that need to traverse the WAN.

Scenario 19-2 Answers

Refer to Figure 19-2 to see the topology for this scenario.

1. *Router A was chosen as the designated router on the LAN in the OSPF domain. Explain why this design choice was made, and give the command that would ensure its selection as the designated router.*

The router was selected by the administrator to be the designated router because it is an Autonomous System Boundary Router (ASBR). This makes it the most sensible choice. All traffic must pass through it to reach another domain. It should also be a more powerful router, because it must calculate redistribution and filtering and handle the role of designated router. The configuration to ensure that this router is chosen as the designated router is as follows:

```
Router(config)#router ospf 100
Router(config)#interface Ethernet0
Router(config-if)#ip ospf priority 100
```

The designated router will have been selected manually by using this priority command or by configuring the router with the highest OSPF ID. This is achieved by allocating a loopback address. The default priority value is 1. If the value is set to 0, the router is ineligible to be a designated router (DR).

2. *Redistribution is occurring between IGRP and OSPF on Router A; state the configuration commands that you would use for both IGRP and OSPF.*

The configuration commands for redistributing OSPF into IGRP and IGRP back into OSPF follow:

```
Router(config)#router igrp 100
Router(config-router)#passive-interface Ethernet0
Router(config-router)#network 201.100.10.0
Router(config-router)#redistribute ospf 100 metric 10000 100 255 1 1500
!
Router(config)#router ospf 100
Router(config-router)#passive-interface e1
Router(config-router)#network 144.250.0.0 0.0.255.255 area 0
Router(config-router)#redistribute igrp 100 metric 30 metric-type 1 subnets
```

The use of the passive interface is unnecessary because the routing processes do not use the same Internet number. These commands will allow full connectivity between the two domains.

The metric has been specified as 30. This is the seed metric that will be used on the routes entering OSPF from EIGRP. Because the routes are external, the route is made to appear farther away than the connected routes. The metric type is identified as type 1 external, meaning that the route will include the internal cost to the ASBR added to the external cost reported by the ASBR.

3. *The new headquarters of the merged company is in the OSPF domain. Therefore, because the connection to the Internet is also in this domain, it has been decided that Router A will have a default route configured to point to the OSPF domain from the IGRP domain. It will be necessary to configure the **ip classless** command for the IGRP domain. Explain why the **ip classless** command is needed, and tell what it achieves.*

Normally, a classful routing protocol tries to match a destination address to a subnet it is attached to or to a remote classful network. Failing to find a match in either case causes the traffic to be dropped. Using the **ip classless** command amends this behavior so that unmatched traffic is also compared to less-specific routes (supernets or the default route) to try to find a match.

4. *Explain the use of the **subnet** command in OSPF, and why it is required in redistribution.*

The **subnet** command is used to propagate subnetworks into the OSPF domain instead of propagating the larger classful address.

5. *In a more complex topology with multiple redistribution points, explain why it might be necessary to filter when redistributing, and what the filter achieves in this network.*

The filter prevents feedback from OSPF into IGRP, and vice versa. In a network with more than one redistribution point, the networks can be redistributed at point A and then redistributed back into the domain that originated the route. Filtering out these routes avoids routing loops.

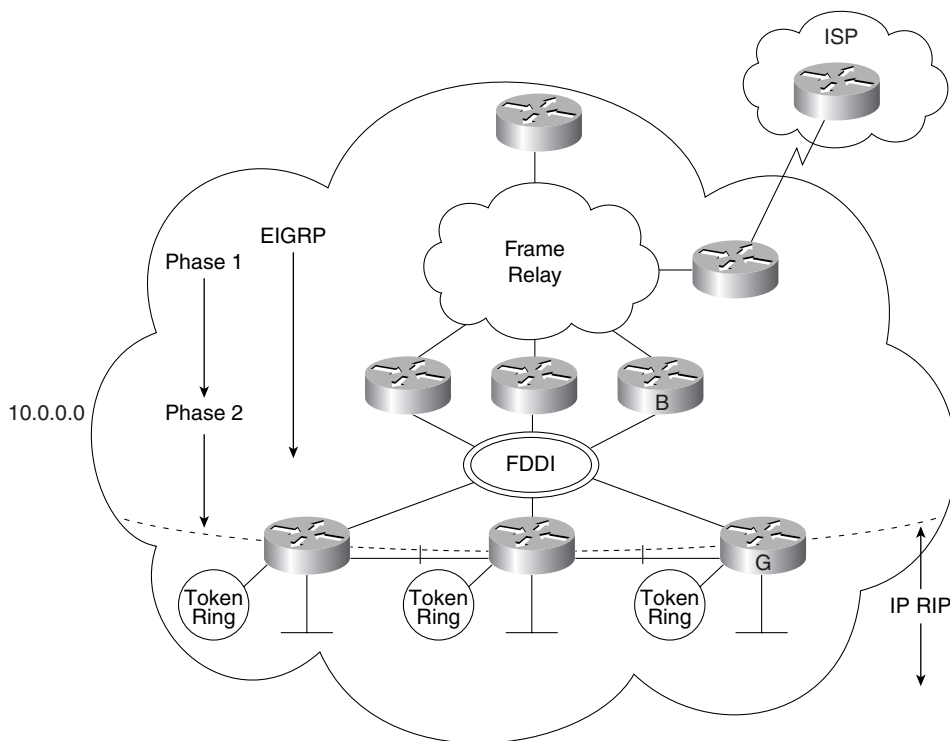
Scenario 19-3 Answers

1. *Create a transition plan, showing how the final design will be implemented.*

Figure 19-4 shows the answer to this question.

In Figure 19-4, the transition would start at the core and move down to the distribution layer routers. The first move would be to place EIGRP in the Frame Relay cloud to alleviate the network overhead. After this is operational, the second phase would be to implement EIGRP in the FDDI ring because there are no end stations on the ring that require IPX RIP/SAP.

Figure 19-4 Answer Diagram for Scenario 19-3, Question 1



2. If the transition plan requires redistribution, indicate these points on the diagram and provide the configuration commands.

The configuration commands for Router G that provide for the redistribution between the different routing protocols are as follows:

```
Router(config)# router rip
Router(config-router)#network 10.0.0.0
Router(config-router)#passive-interface Ethernet 0
Router(config-router)#passive interface Token 0
Router(config-router)#redistribute eigrp 100
Router(config-router)# default-metric 3
Router(config)# router eigrp 100
Router(config-router)#network 140.100.0.0
Router(config-router)#passive-interface FDDI 0
Router(config-router)#redistribute rip
Router(config-router)# default-metric 10000 100 255 1 1500
```

3. The company has also decided to connect to the Internet using EIGRP, upon which the ISP that is providing the Internet connectivity has agreed. Is any redistribution required? Give reasons for your answers.

If the ISP is in the same autonomous system as the company, no redistribution is required between the company and the ISP. Within the ISP domain, however, it is likely that the router connecting to the company in question is the only router that is a part of the autonomous system of the company. This is for security reasons. Therefore, some form of redistribution probably will be required in the ISP domain.

4. *Explain the neighbors that you see on the LAN segment and the path selection on Router B to network 140.100.0.0.*

The neighbors that are seen on the FDDI ring are the six routers that share the physical medium. Router B chooses the path as a successor because it provided the shortest path to the network stated, that is, a direct route through H.

5. *Create a diagram that shows the final network design, with configuration commands for EIGRP on Router A.*

Figure 19-5 shows the answer.

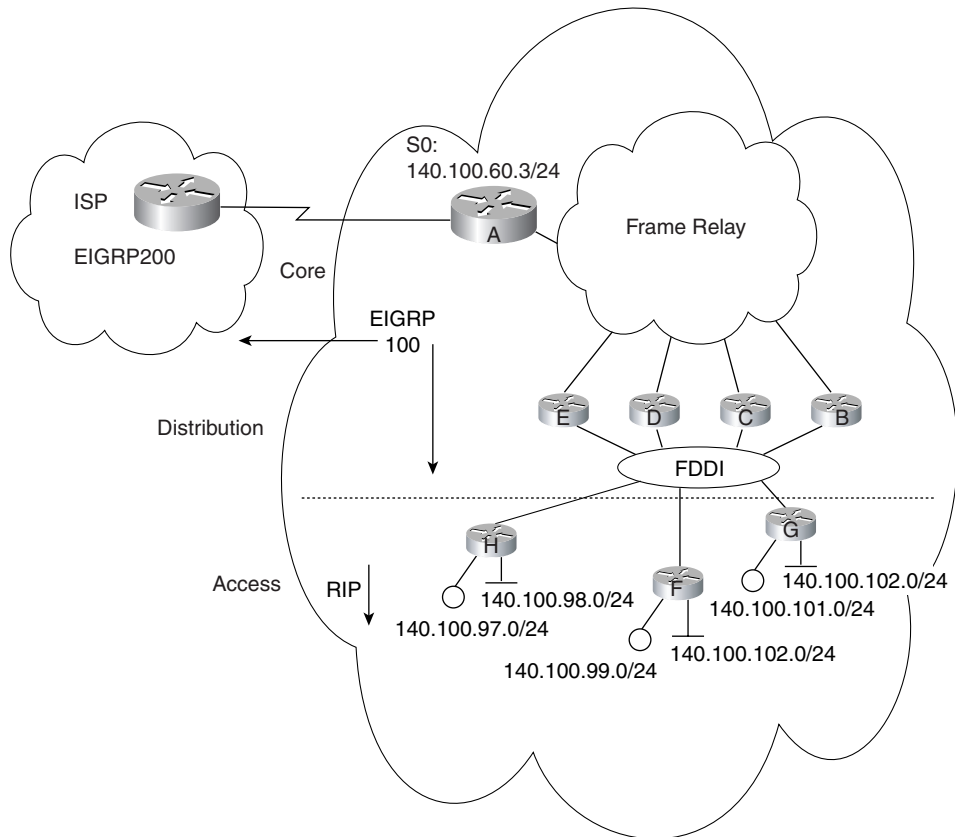
In Figure 19-5, the command **ip summary-address** can be configured where appropriate at the distribution layer (the routers entering the Frame Relay cloud):

```
Router A
Router(config)#router EIGRP 100
Router(config-router)#network 140.100.0.0
Router(config-router)#ip summary-address eigrp 140.100.0.0 255.255.0.0
```

6. *Write out the configuration commands for Router B.*

The configuration commands for Router B are as follows:

```
Router(config)#router eigrp 100
Router(config-router)#network 140.100.0.0
Router(config-router)#no auto-summary
Router(config-router)#ip summary-address 140.100.64.0 255.255.224.0
```

Figure 19-5 Answer Diagram to Scenario 19-3, Question 5

Three-Part Scenarios

This section contains three three-part scenarios that require planning, configuration, and verification. The solutions to these three-part scenarios are contained within each scenario.

Scenario 19-4

Part A of Scenario 19-4 begins with some planning guidelines that include planning IP addresses, designing the VLSM addressing scheme, identifying the OSPF areas, and determining what type of areas they should be. After you complete Part A, Part B of the scenario asks you to configure the three routers to implement the planned design and a few other features. Finally, Part C asks you to examine router command output to discover details about the current operation. Part C also lists some questions related to the user interface and protocol specifications.

Scenario 19-4, Part A—Planning

Your job is to deploy a new network with three sites, as shown in Figure 19-6.

The OSPF network has a shortage of IP addresses. It has been decided to readdress the network using VLSM. For Part A of this scenario, perform the following tasks:

1. Plan the IP addressing, using the Class B address of 131.99.0.0. Each site consists of two buildings, with seven floors. Each floor has approximately 100 devices. The company plans to install an ISDN backup link between the buildings.
2. In the expectation of growth, the company has decided that each site should be an area. Currently, each site has only two buildings. The floors of each building are connected via a switch. One of the sites has a department running UNIX servers that are using RIP. The RIP networks are redistributed into the OSPF network.

Define the location of the areas.

For each router, define the router type and the number of each area connected. State your reasons for your choices.

Figure 19-6 *Scenario 19-4 Network Diagram*

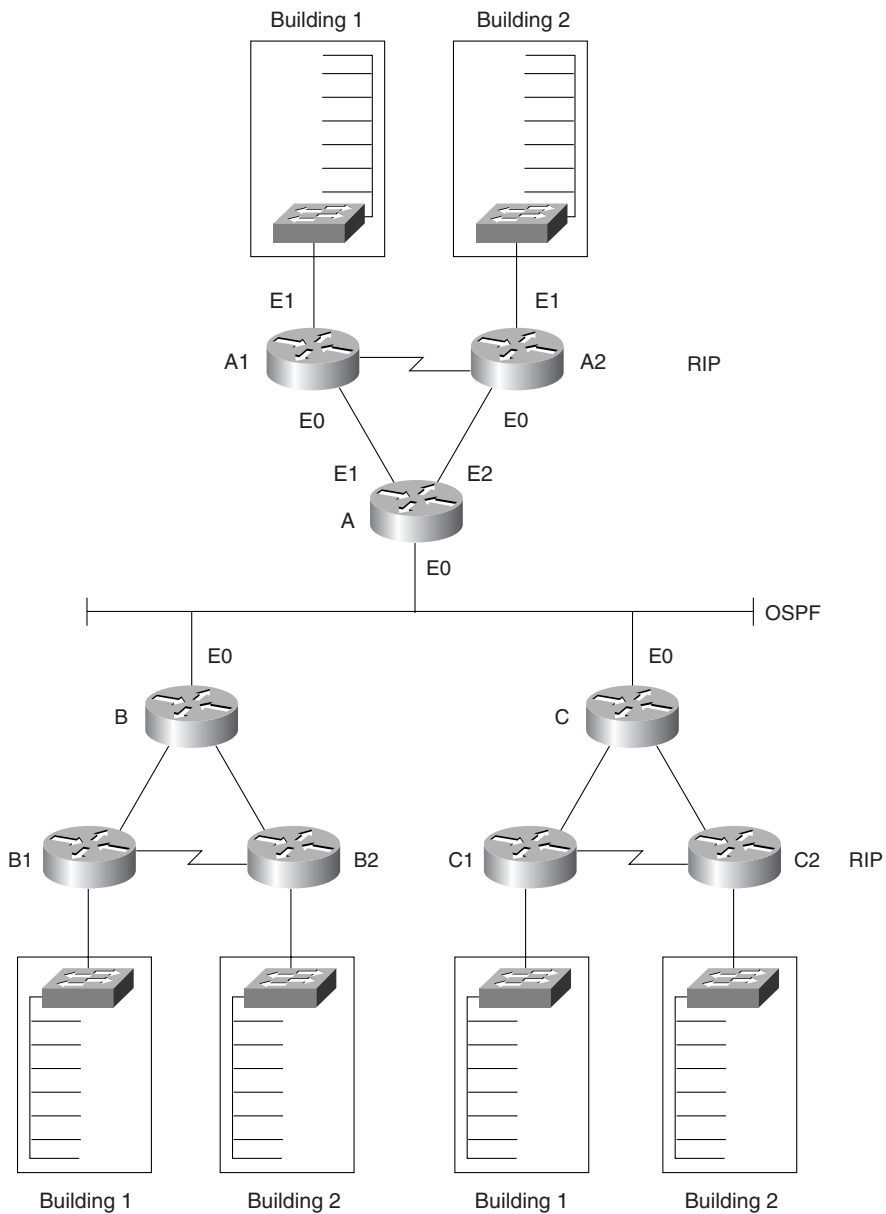


Table 19-1 and Table 19-2 are provided to show how to record your IP subnets when performing the planning tasks for this scenario.

Table 19-1 Scenario 19-4, Part A—IP Subnet Network Planning Chart

Geographic Location of Subnet/Network	Subnet Number	Bit Pattern of Last Two Octets	Prefix	Subnet Mask
A Ethernet 0				
B Ethernet 0				
C Ethernet 0				
Serial between A1 and A2				
A1 Ethernet 1 VLANs Floor 1 Floor 2 Floor 3 Floor 4 Floor 5 Floor 6 Floor 7				

Table 19-1 Scenario 19-4, Part A—IP Subnet Network Planning Chart (Continued)

Geographic Location of Subnet/Network	Subnet Number	Bit Pattern of Last Two Octets	Prefix	Subnet Mask
A2				
Ethernet 1				
VLANs				
Floor 1				
Floor 2				
Floor 3				
Floor 4				
Floor 5				
Floor 6				
Floor 7				
Serial between B1 and B2				
B1				
Ethernet 1				
VLANs				
Floor 1				
Floor 2				
Floor 3				
Floor 4				
Floor 5				
Floor 6				
Floor 7				

(continues)

Table 19-1 Scenario 19-4, Part A—IP Subnet Network Planning Chart (Continued)

Geographic Location of Subnet/Network	Subnet Number	Bit Pattern of Last Two Octets	Prefix	Subnet Mask
B2				
Ethernet 1				
VLANs				
Floor 1				
Floor 2				
Floor 3				
Floor 4				
Floor 5				
Floor 6				
Floor 7				
Serial between C1 and C2				
C1				
Ethernet 1				
VLANs				
Floor 1				
Floor 2				
Floor 3				
Floor 4				
Floor 5				
Floor 6				
Floor 7				

Table 19-1 Scenario 19-4, Part A—IP Subnet Network Planning Chart (Continued)

Geographic Location of Subnet/Network	Subnet Number	Bit Pattern of Last Two Octets	Prefix	Subnet Mask
C2				
Ethernet 1				
VLANs				
Floor 1				
Floor 2				
Floor 3				
Floor 4				
Floor 5				
Floor 6				
Floor 7				

Table 19-2 Scenario 19-4, Part A—OSPF Area Planning Chart

Router Location	Router Type	Reason
A		
B		
C		
A1		
A2		
B1		
B2		
C1		
C2		

Solutions to Scenario 19-4, Part A—Planning

Keeping the design as simple as possible—yet not too simple, so that it is still useful as the network evolves—is a good practice. In the suggested answers in Tables 19-3 and 19-4, a numbering scheme is presented. Remember, this is one of many solutions available. The reasoning behind this planning is to allow summarization at the area border routers (ABRs). The first few bits in the third octet indicate the summarization bits. Note that these bits reflect the area; 1 bit, or 128, identifies Area 1, while 3 bits, or 224, identify Area 3.

Table 19-3 Scenario 19-4, Part A—IP Subnet Network Planning Chart Solution

Geographic Location of Subnet/Network	Subnet Number	Bit Pattern of Last Two Octets	Prefix	Subnet Mask
A				
Ethernet 0	131.99.0.48	00000000 0011 0000	/28	255.255.255.240
Ethernet 1	131.99.136.128	10001000 1000 0000	/28	255.255.255.240
Ethernet 2	131.99.144.128	10010000 1000 0000	/28	255.255.255.240
B				
Ethernet 0	131.99.0.48	00000000 0011 0000	/28	255.255.255.240
Ethernet 1	131.99.192.192	11000000 1100 0000	/28	255.255.255.240
Ethernet 2	131.99.216.192	11011000 1100 0000	/28	255.255.255.240
C				
Ethernet 0	131.99.0.48	00000000 0011 0000	/28	255.255.255.240
Ethernet 1	131.99.224.224	11100000 1110 0000	/28	255.255.255.240
Ethernet 2	131.99.232.224	11101000 1110 0000	/28	255.255.255.240
Serial between A1 and A2	131.99.128.128	10000000 1000 0000	/26	255.255.255.192

Table 19-3 *Scenario 19-4, Part A—IP Subnet Network Planning Chart Solution (Continued)*

Geographic Location of Subnet/Network	Subnet Number	Bit Pattern of Last Two Octets	Prefix	Subnet Mask
A1				
Ethernet 0	131.99.136.128	10001000 1000 0000	/28	255.255.255.240
Ethernet 1				
VLANs				
Floor 1	131.99.136.0	10001000 0 0000000	/25	255.255.255.128
Floor 2	131.99.137.0	10001001 0 0000000	/25	255.255.255.128
Floor 3	131.99.138.0	10001010 0 0000000	/25	255.255.255.128
Floor 4	131.99.139.0	10001011 0 0000000	/25	255.255.255.128
Floor 5	131.99.140.0	10001100 0 0000000	/25	255.255.255.128
Floor 6	131.99.141.0	10001101 0 0000000	/25	255.255.255.128
Floor 7	131.99.142.0	10001110 0 0000000	/25	255.255.255.128
A2				
Ethernet 0	131.99.144.128	10010000 1000 0000	/28	255.255.255.240
Ethernet 1				
VLANs				
Floor 1	131.99.144.0	10010000 0 0000000	/25	255.255.255.128
Floor 2	131.99.145.0	10010001 0 0000000	/25	255.255.255.128
Floor 3	131.99.146.0	10010010 0 0000000	/25	255.255.255.128
Floor 4	131.99.147.0	10010011 0 0000000	/25	255.255.255.128
Floor 5	131.99.148.0	10010100 0 0000000	/25	255.255.255.128
Floor 6	131.99.149.0	10010101 0 0000000	/25	255.255.255.128
Floor 7	131.99.150.0	10010110 0 0000000	/25	255.255.255.128

(continues)

Table 19-3 *Scenario 19-4, Part A—IP Subnet Network Planning Chart Solution (Continued)*

Geographic Location of Subnet/Network	Subnet Number	Bit Pattern of Last Two Octets	Prefix	Subnet Mask
Serial between B1 and B2	131.99.192.128	11000000 10 000000	/26	255.255.255.192
B1				
Ethernet 0	131.99.192.192	11000000 1100 0000	/28	225.255.255.240
Ethernet 1				
VLANs				
Floor 1	131.99.200.0	11001000 0 0000000	/25	255.255.255.128
Floor 2	131.99.201.0	11001001 0 0000000	/25	255.255.255.128
Floor 3	131.99.202.0	11001010 0 0000000	/25	255.255.255.128
Floor 4	131.99.203.0	11001011 0 0000000	/25	255.255.255.128
Floor 5	131.99.204.0	11001100 0 0000000	/25	255.255.255.128
Floor 6	131.99.205.0	11001101 0 0000000	/25	255.255.255.128
Floor 7	131.99.206.0	11001110 0 0000000	/25	255.255.255.128
B2				
Ethernet 0	131.99.216.192	11011000 1100 0000	/28	225.255.255.240
Ethernet 1				
VLANs				
Floor 1	131.99.216.0	11011000 0 0000000	/25	255.255.255.128
Floor 2	131.99.217.0	11011001 0 0000000	/25	255.255.255.128
Floor 3	131.99.218.0	11011010 0 0000000	/25	255.255.255.128
Floor 4	131.99.219.0	11011011 0 0000000	/25	255.255.255.128
Floor 5	131.99.220.0	11011100 0 0000000	/25	255.255.255.128
Floor 6	131.99.221.0	11011101 0 0000000	/25	255.255.255.128
Floor 7	131.99.222.0	11011110 0 0000000	/25	255.255.255.128

Table 19-3 *Scenario 19-4, Part A—IP Subnet Network Planning Chart Solution (Continued)*

Geographic Location of Subnet/Network	Subnet Number	Bit Pattern of Last Two Octets	Prefix	Subnet Mask
Serial between C1 and C2	131.99.224.128	11100000 10 000000	/26	255.255.255.192
C1				
Ethernet 0	131.99.224.224	11100000 1110 0000	/28	225.255.255.240
Ethernet 1				
VLANs				
Floor 1	131.99.224.0	11100000 0 0000000	/25	255.255.255.128
Floor 2	131.99.225.0	11100001 0 0000000	/25	255.255.255.128
Floor 3	131.99.226.0	11100010 0 0000000	/25	255.255.255.128
Floor 4	131.99.227.0	11100011 0 0000000	/25	255.255.255.128
Floor 5	131.99.228.0	11100100 0 0000000	/25	255.255.255.128
Floor 6	131.99.229.0	11100101 0 0000000	/25	255.255.255.128
Floor 7	131.99.230.0	11100110 0 0000000	/25	255.255.255.128
C2				
Ethernet 0	131.99.232.224	11101000 1110 0000	/28	255.255.255.240
Ethernet 1				
VLANs				
Floor 1	131.99.232.0	11101000 0 0000000	/25	255.255.255.128
Floor 2	131.99.233.0	11101001 0 0000000	/25	255.255.255.128
Floor 3	131.99.234.0	11101010 0 0000000	/25	255.255.255.128
Floor 4	131.99.235.0	11101011 0 0000000	/25	255.255.255.128
Floor 5	131.99.236.0	11101100 0 0000000	/25	255.255.255.128
Floor 6	131.99.237.0	11101101 0 0000000	/25	255.255.255.128
Floor 7	131.99.238.0	11101110 0 0000000	/25	255.255.255.128

Table 19-4 *Part A—OSPF Area Planning Chart Solution*

Router Location	OSPF Router Type	Reason
A in Area 0 and Area 1	Backbone router and ABR. The interface into Area 1 is configured as a stub area.	The router has an interface in both Area 0 and Area 1; it is at the core of the network. The interface into Area 1 will not send external routes into the area, but it might generate default route and summary routes.
B in Area 0 and Area 2	Backbone router and ABR. The interface into Area 2 is configured as a stub area.	The router has an interface in both Area 0 and Area 2; it is at the core of the network. The interface into Area 2 will not send external routes into the area, but it might generate default route and summary routes.
C in Area 0 and Area 3	Backbone router and ABR. The interface into Area 3 is <i>not</i> configured as a stub area because the area has RIP redistributed into the area.	The router has an interface in both Area 0 and Area 3; it is at the core of the network. The interface into Area 3 will not send external routes into the area, but the router generates a default route into Area 3. Because this area is the only area that has external routes generated within the area through redistribution from RIP, this is not an issue. However, if the organization ever connects to the Internet, additional configuration will be necessary to distinguish between the Internet and the RIP network.
A1	Internal router and stub area router into Area 1.	It is in Area 1 only using summarization, with no need for the RIP networks; it can therefore be configured as a stub area.
A2	Internal router and stub area router into Area 1.	It is in Area 1 only using summarization with no need for the RIP networks; it can therefore be configured as a stub area.
B1	Internal router and stub area router into Area 2.	It is in Area 2 only using summarization with no need for the RIP networks; it can therefore be configured as a stub area.
B2	Internal router and stub area router into Area 2.	It is in Area 2 only using summarization with no need for the RIP networks; it can therefore be configured as a stub area.
C1	Internal router to Area 3.	This router cannot be a stub area router because the area has RIP redistributed into it.
C2	An autonomous system border router (ASBR)	Because the building that this router serves has UNIX servers running RIP, and because these networks need to be accessible to the organization, RIP needs to be redistributed into the OSPF domain. Redistributed routes are considered to be external routes—thus the configuration of an ASBR.

Scenario 19-4, Part B—Configuration

The next step in your job is to deploy the network designed in Scenario 19-4, Part A. Use the solutions for Part A of Scenario 19-4 to direct you in configuring the addressing and summarization at the ABRs. For Scenario 19-4, Part B, perform the following tasks:

1. Configure basic OSPF for Routers A, B, and C based on the design in Scenario 19-4, Part A.
2. Configure summarization on Routers A, B, and C.
3. Configure Routers A and B to connect to stub areas.

Solutions to Scenario 19-4, Part B—Configuration

Example 19-1, Example 19-2, and Example 19-3 show the configurations for Scenario 19-4, Part B, given the criteria.

Example 19-1 Scenario 19-4 Router A Configuration for Questions 1, 2, and 3

```
Router(config)#interface Ethernet0
Router(config-if)#ip address 131.99.0.49 255.255.255.240
! 4 bits allows 14 hosts
Router(config)#interface e1
Router(config-if)#ip address 131.99.136.129 255.255.255.240
! 4 bits allows 14 hosts
Router(config)#interface e2
Router(config-if)#ip address 131.99.144.129 255.255.255.240
! 4 bits allows 14 hosts
!
Router(config)#router ospf 100
Router(config-router)#network 131.99.0.49 0.0.0.0 area 0
! The wildcard mask 0.0.0.0 places the interface ETHERNET0 into area 0
Router(config-router)#network 0.0.0.0 255.255.255.255 area 1
! The wildcard mask 255.255.255.255 places all other interfaces into area 1

!!Answer to Question 2 for Router A--Configure Summarization
! All addresses from area 1 will be summarized into this one network
Router(config-router)#area 1 range 131.99.128.0 255.255.224.0
!!Answer to Question 3 for Router A--Configure Routers A and B to connect to stub areas
[Frame1]
```

Example 19-2 *Scenario 19-4 Router B Configuration for Questions 1, 2, and 3*

```

Router(config)#interface Ethernet0
Router(config-if)#ip address 131.99.0.50 255.255.255.240
! 4 bits allows 14 hosts
Router(config)#interface e1
Router(config-if)#ip address 131.99.192.193 255.255.255.240
! 4 bits allows 14 hosts
Router(config)#interface e2
Router(config-if)#ip address 131.99.216.193 255.255.255.240
! 4 bits allows 14 hosts
!!
Router(config)#router ospf 100
Router(config-router)#network 131.99.0.50 0.0.0.0 area 0
Router(config-router)#network 0.0.0.0 255.255.255.255 area 2
!!Answer to Question 2 for Router B--Configure Summarization
! All addresses from area 2 will be summarized into this one network
Router(config-router)#area 2 range 131.99.192.0 255.255.224.0
!!Answer to Question 3 for Router B to connect to stub areas and to summarize routes
[Frame2]
Router(config-router)#area 2 stub

```

Example 19-3 *Scenario 19-4 Router C Configuration for Questions 1, 2, and 3*

```

Router(config)#interface Ethernet0
Router(config-if)#ip address 131.99.0.51 255.255.255.240
! 4 bits allows 14 hosts
Router(config)#interface e1
Router(config-if)#ip address 131.99.224.225 255.255.255.128
! 7 bits allows 127 hosts
Router(config)#interface e2
Router(config-if)#ip address 131.99.232.225 255.255.255.128
! 4 bits allows 14 hosts
!!
Router(config)#router ospf 100
Router(config-router)#network 131.99.0.51 0.0.0.0 area 0
router(config-router)#network 0.0.0.0 255.255.255.255 area 3
! All addresses from area 3 will be summarized into this one network
!!Answer to Question 2 for Router C--Configure Summarization
Router(config-router)#area 3 range 131.99.224.0 255.255.224.0

```

Scenario 19-4, Part C—Verification and Questions

Answer the questions following Example 19-4. Use Example 19-4 as a reference when the question refers directly to this scenario. Although not all of these questions are directly tied to Part A of this scenario, they all probe foundational knowledge required by the technology examined in this scenario.

NOTE In the network from which these commands were captured, several administrative settings not mentioned in the scenario were configured. For instance, the **enable** password was configured. Any **show running-config** commands in the examples in this chapter might have other unrelated configuration.

Example 19-4 *The show and debug Output Screens for Scenario 19-4, Router A*

```
RouterA#show ip protocols
Routing Protocol is "ospf 100"
  Sending updates every 0 seconds
  Invalid after 0 seconds, hold down 0, flushed after 0
  Outgoing update filter list for all interfaces is
  Incoming update filter list for all interfaces is
  Redistributing: ospf 100
  Routing for Networks:
    131.99.0.49/32
    0.0.0.0
  Routing Information Sources:
    Gateway          Distance      Last Update
    131.99.232.225    110          00:14:49
    131.99.216.193    110          00:14:49
  Distance: (default is 110)

RouterA# show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, ia - IS-IS inter area
       * - candidate default, U - per-user static route, o - ODR
       P - periodic downloaded static route

Gateway of last resort is not set

    131.99.0.0/16 is variably subnetted, 6 subnets, 4 masks
O IA   131.99.224.0/20 [110/11] via 131.99.0.51, 00:14:56, Ethernet0
O IA   131.99.192.0/20 [110/11] via 131.99.0.50, 00:14:56, Ethernet0
C       131.99.0.48/28 is directly connected, Ethernet0
C       131.99.136.128/25 is directly connected, Ethernet1

RouterA#show ip ospf database
        OSPF Router with ID (131.99.144.129) (Process ID 100)

        Router Link States (Area 0)
```

continues

Example 19-4 The **show** and **debug** Output Screens for Scenario 19-4, Router A (Continued)

Link ID	ADV Router	Age	Seq#	Checksum	Link count
131.99.144.129	131.99.144.129	1071	0x80000004	0xEB16	1
131.99.216.193	131.99.216.193	1122	0x80000004	0xETHERNET00F	1
131.99.232.225	131.99.232.225	1122	0x80000004	0x296E	1
network Link States (Area 0)					
Link ID	ADV Router	Age	Seq#	Checksum	
131.99.0.51	131.99.232.225	1072	0x80000002	0x516B	
Summary network Link States (Area 0)					
Link ID	ADV Router	Age	Seq#	Checksum	
131.99.128.0	131.99.144.129	1079	0x80000001	0xB436	
131.99.144.129	131.99.144.129	1079	0x80000001	0x400A	
131.99.192.0	131.99.216.193	1129	0x80000001	0x76AB	
131.99.216.193	131.99.216.193	1129	0x80000001	0x2712	
131.99.224.0	131.99.232.225	1215	0x80000001	0xE3ED	
Router Link States (Area 1)					
Link ID	ADV Router	Age	Seq#	Checksum	Link count
131.99.144.129	131.99.144.129	1080	0x80000003	0x53F	2
Summary network Link States (Area 1)					
Link ID	ADV Router	Age	Seq#	Checksum	
0.0.0.0	131.99.144.129	1080	0x80000001	0xF94B	
131.99.0.48	131.99.144.129	1070	0x80000001	0x1D3	
131.99.192.0	131.99.144.129	1070	0x80000001	0x742E	
131.99.216.193	131.99.144.129	1071	0x80000001	0x2594	
RouterA#show ip ospf border routers					
OSPF Process 100 internal Routing Table					
Codes: i - Intra-area route, I - Inter-area route					
i 131.99.232.225 [10] via 131.99.0.51, Ethernet0, ABR, Area 0, SPF 6					
i 131.99.216.193 [10] via 131.99.0.50, Ethernet0, ABR, Area 0, SPF 6					
RouterA#show ip ospf neighbor					
Neighbor ID	Pri	State	Dead Time	Address	Interface
131.99.232.225	1	FULL/DR	00:00:39	131.99.0.51	Ethernet0
131.99.216.193	1	FULL/BDR	00:00:38	131.99.0.50	Ethernet0

Answer the following questions:

1. Which command do you use to configure a totally stubby area?
2. What do the letters *ASBR* stand for, and what does this device do?
3. Where is summarization performed in OSPF?
4. Give the command to configure the ASBR to summarize the networks 131.99.224.128, 131.99.224.224, and 131.99.224.0 through 131.99.230.0 for redistribution into the RIP process, using a 20-bit subnet mask.
5. Explain the difference between prefix routing and subnetting.
6. State one consideration when configuring multiarea OSPF across a nonbroadcast multiaccess (NBMA) network.
7. What command do you use to turn on OSPF?
8. Explain why OSPF supports VLSM.
9. Explain why all areas must connect through the backbone Area 0.
10. Explain the purpose of the **network** command in OSPF.
11. Which command do you use to show a router's internal OSPF routing table?
12. Which command shows the use of VLSM on the OSPF network?
13. Which command verifies the establishment of adjacencies with other routers on the same network?
14. How would you determine whether there is more than one IP routing protocol running on a router? If more than one IP routing protocol is running, how would you know how redistribution was configured?
15. When troubleshooting an OSPF configuration over an NBMA network, which command shows the network type that has been configured? The command **show running** will display this information, but greater analysis of the timers and costs is required.
16. Give the appropriate mask to use on a point-to-point serial interface, where **IP unnumbered** is not an option.
17. What command would you use to identify that an adjacency could not be formed because one router was configured as a stub, while another was not?
18. Which command do you use to ensure that the virtual link is active?
19. What concern should you have when using the **debug** command?
20. What is a floating static route, and when would you use one?

Solutions to Scenario 19-4, Part C—Verification and Questions

The answers to the questions for Scenario 19-4, Part C, are as follows:

1. *Which command do you use to configure a totally stubby area?*

The command used to create a totally stubby area is as follows:

```
Router(config-router)# area area-id stub no-summary
```

2. *What do the letters ASBR stand for, and what does this device do?*

The letters *ASBR* stand for autonomous system boundary router. This is an OSPF router that connects the OSPF domain to another routing domain. This is when redistribution is required. The ASBR does not have to be situated in Area 0, although that is recommended.

3. *Where is summarization performed in OSPF?*

Summarization is configured on the ABR or the ASBR.

Summarization at the ABR creates an advertisement that represents many networks within the area. This summary advertisement is propagated into Area 0 and, from there, into the other areas.

Summarization at the ASBR creates an advertisement that represents many networks within the autonomous system. Redistributing this summary advertisement into another routing protocol propagates it to the outside world.

4. *Give the command to configure the ASBR to summarize the networks 131.99.224.128, 131.99.224.224, and 131.99.224.0 through 131.99.230.0 for redistribution into the RIP process, using a 20-bit subnet mask.*

The command to summarize the networks 131.99.224.128, 131.99.224.224, and 131.99.224.0 through 131.99.240.0 for redistribution into the RIP process is as follows:

```
Router(config-router)#summary-address 131.99.224.0 255.255.240.0
```

Because RIPv1 is using the same classful network address, it is important to note that RIP would need to use the same mask because it does not support VLSM.

5. *Explain the difference between prefix routing and subnetting.*

The difference between prefix routing and subnetting is that prefix routing creates supernets for the Internet. That summarizes class addresses, for example, creating one network from 16 Class C addresses. The new mask would be 255.255.240.0 or a prefix of /20.

Subnetting is the means by which the classful address is broken down to create many smaller networks addressing fewer hosts, instead of one network with the capability to address many hosts. VLSM is the capability to subnet the class address provided by the Internet, to create

subnets with different subnet masks, so that the number of hosts can vary according to the need of the segment.

Prefix routing moves the mask to the left, whereas subnetting moves the mask to the right.

6. *State one consideration when configuring multiarea OSPF across a nonbroadcast multiaccess (NBMA) network.*

Considerations when configuring multiarea OSPF across a nonbroadcast multiaccess (NBMA) network include these:

- The NBMA network can be created as Area 0. The reasoning is that if the NBMA is used to connect all remote sites, all traffic will have to traverse this network. If the remote sites are satellite areas, all traffic would have to traverse the NBMA, so it makes sense to make Area 0 the backbone area. This works well in a full-mesh environment, although it results in a large number of LSAs being flooded into the WAN and puts extra demands on the routers connecting to the NBMA network.
- In a hub-and-spoke NBMA network, it makes sense to assign the hub network as Area 0 with the other remote sites and the NBMA network as other areas. This is a good design if the satellite areas are stub areas, because it means that the routing information, and thus network overhead, is kept to a minimum over the NBMA cloud. Depending on the design, the rest of the network might constitute one other area or multiple areas. This will depend on the size and growth expectations of the OSPF domain.

7. *What command do you use to turn on OSPF?*

The command to turn on OSPF is as follows:

```
Router(config)# Router OSPF process-id
```

The *process-id* is subtly different from the autonomous system number used in IGRP and EIGRP. The *process-id* identifies the OSPF routing process on the router and has no significance outside the router. This allows more than one process to be configured on a router. Although it is unusual, there are instances in which this configuration is appropriate.

8. *Explain why OSPF supports VLSM.*

OSPF supports VLSM because it carries the subnet mask in the routing updates. Therefore, each router can reference the appropriate mask for each network.

9. *Explain why all areas must connect through the backbone Area 0.*

Area 0 forms a common path for all areas to connect. Therefore, it ensures that all areas are aware of all networks within the OSPF domain.

10. *Explain the purpose of the **network** command in OSPF.*

The **network** command is used to assign an interface or a group of interfaces to an area. When the interfaces are identified, they will participate in the OSPF routing process for the area to which they belong.

11. *Which command do you use to show a router's internal OSPF routing table?*

The command **show ip ospf border-routers** displays the internal routing table of the OSPF internal router.

12. *Which command shows the use of VLSM on the OSPF network?*

The **show ip route** command shows not only all the available routes, but also the masks used. If different masks are used, this command states that the subnet is variably subnetted, with the number of subnets and masks.

13. *Which command verifies the establishment of adjacencies with other routers on the same network?*

The command **show ip ospf neighbor** shows the neighbors, the neighbor ID, the connecting interface address of the neighboring router, the outgoing interface on the router to connect to the neighbor, the length of time since the last communication from the neighbor, the connection state, and whether the neighbor is a designated router (DR) or backup designated router (BDR).

14. *How would you determine whether there is more than one IP routing protocol running on a router? If more than one IP routing protocol is running on a router, how would you know how redistribution is configured?*

There are a couple of ways to ascertain whether more than one routing protocol is running on a router. The first way is by issuing this command:

```
Router# show running
```

Because this command shows the configuration that is currently running, it is easy enough to see if more than one protocol is running and, if redistribution is running, how it is configured. The other and more straightforward method is to issue the following command:

```
Router# show ip protocols
```

This command shows all the IP protocols configured on the router, as well as the details of that configuration, including the redistribution.

15. *When troubleshooting an OSPF configuration over an NBMA network, which command shows the network type that has been configured? The command **show running** will display this information, but greater analysis of the timers and costs is required.*

The command **show ip ospf interface** indicates the network type, the delays set for the timers, and the number of neighbors and adjacencies.

16. *Give the appropriate mask to use on a point-to-point serial interface, where **IP unnumbered** is not an option.*

The mask to be used in a point-to-point link is 255.255.255.252. This mask allows the allocation of two addresses and is a typical mask for point-to-point networks.

17. *What command would you use to identify that an adjacency could not be formed because one router was configured as a stub, while another was not?*

The command **debug ip ospf adjacency** will show that there is a problem in the establishment of an adjacency. The error message would point to a mismatched stub/transit area option bit. You could also do a **show ip ospf interface** to see that the interface has no neighbors. At this point, it would be wise to check the interface configuration.

18. *Which command do you use to ensure that the virtual link is active?*

The command **show ip ospf virtual-links** verifies the virtual link configuration.

19. *What concern should you have when using the **debug** command?*

The **debug** command is a command that should be used with caution because it can use an enormous amount of system resources. Because it is given priority 1 as a process, it can eventually bring your router to a standstill. It is important to run the utility for a limited time to capture the required output. It is also best not to use the **debug** command from the console unless logging to the log file only. If the console screen freezes while using **debug**, there is no recovery other than to reboot the system.

20. *What is a floating static route, and when would you use one?*

You would use a floating static route when a router has a link that is used as a backup using a medium such as a dialup line. The intention is to have no routing protocol running across the link that would keep the link active at a high cost.

The first task is to create a static route so that the routing protocol is not required. Unfortunately, the administrative distance that is used to select the best routes offered by different routing protocols states that a static route is the best route and will use it for all traffic. Thus, despite your best efforts, the backup link becomes a low-bandwidth, expensive primary link. To change the link to take backup status, manually configure the administrative distance to have a higher value than the dynamic routing protocol.

Scenario 19-5

Part A of Scenario 19-5 begins with some guidelines that include planning the transition of the network to run EIGRP as the routing protocol. To transition smoothly, the intention is to slowly integrate EIGRP; this requires redistribution between EIGRP and OSPF and the filtering of updates between the protocols. After you complete Part A, Part B of the scenario asks you to configure the three routers to implement the planned design and a few other features. Finally, Part C asks you to

examine router command output to discover details about the current operation. Part C also lists some questions related to the user interface and protocol specifications.

Scenario 19-5, Part A—Planning

Your job is to deploy a new network with three sites, as shown in Figure 19-7.

The OSPF network has limitations, and the decision has been made to change the routing protocol to EIGRP. Use the addressing scheme presented in the solutions for Part A of Scenario 19-4. For Part A of this scenario, perform the following tasks:

1. The planning committee has stated that the transition to EIGRP should start at the edges of the network and move inward in a controlled manner. Give a brief explanation of how you would implement a transition plan. Is summarization possible?
2. Redistribution is necessary. State the routers that would be responsible for redistribution and the nature of the redistribution.
3. To prevent routing loops, the company has decided to implement distribution filters. Plan the content, location, and implementation of the filters.

Solutions to Scenario 19-5, Part A—Planning

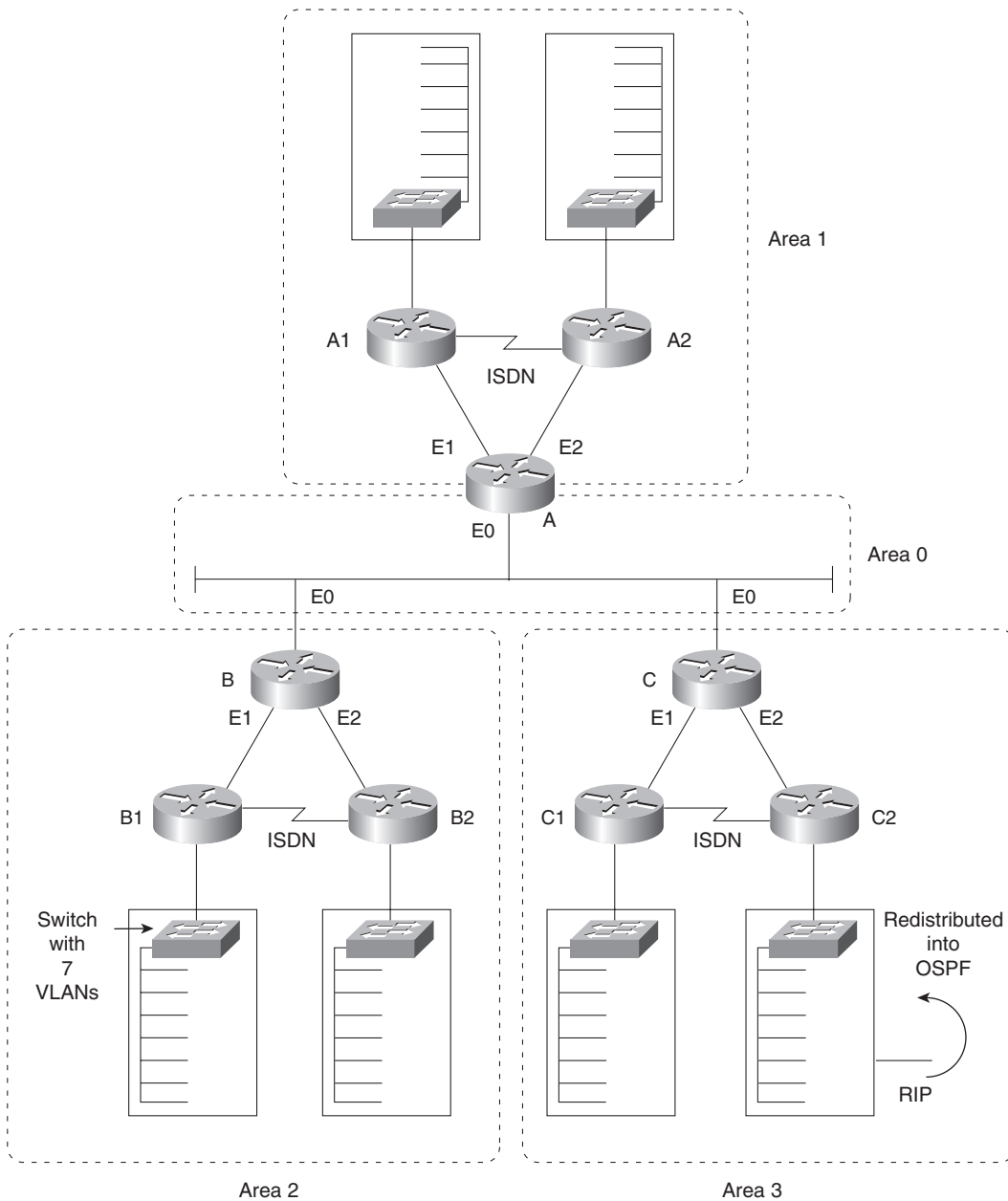
Keeping the design as simple as possible, yet not making it so simple that the network cannot evolve, is a good practice. In these suggested answers, remember that many solutions are available. The reasoning behind this planning is to allow the optimum solution while maintaining the strength of the existing network. As in the Scenario 19-4 solution, the addressing scheme allows summarization at the core level of the network. The first few bits in the third octet indicate the summarization bits.

Many organizations decide to transition the network by starting at the outer peripheries of the network to minimize the impact of these changes. The problems that might occur when a change is made to an organization can be catastrophic. Inevitably, something is overlooked, or you have the dubious honor of discovering a new bug. Obviously, if the change is made to a stub network, the problem can be isolated and fixed speedily. Centralized routers with problems tend to share them with the rest of the network.

1. *The planning committee has stated that the transition to EIGRP should start at the edges of the network and move inward in a controlled manner. Give a brief explanation of how you would implement a transition plan. Is summarization possible?*

The transition plan is straightforward because the network in this scenario is very small. However, it reflects the strategy required by any size network.

Figure 19-7 Scenario 19-5 Network Diagram



The implementation of EIGRP should start in Area 1. The reason for this is that there is currently no redistribution in the area, and it has an addressing scheme that is already set up for summarization.

Summarization is definitely possible. Although EIGRP has no restrictions on summarization, OSPF only permits it at the ABR or ASBR. In EIGRP, summarization is configured at the interface level, which allows for great flexibility, as well as confusion.

There are two ways to consider summarization in this network. The first is simply to implement the summarization put in place by OSPF. The addressing scheme supports it, and it makes good design sense. The other is necessary during the transition phase. This summarization occurs at the point of redistribution, to the external routes being passed into the new routing protocol.

2. *Redistribution is necessary. State the routers that would be responsible for redistribution and the nature of the redistribution.*

Router A would be responsible for redistribution at the first level of transition, which would redistribute the routes from A1. If this transition goes smoothly, A2 would be configured for EIGRP. Then Area 2 could be changed to run EIGRP, one router at a time. The most complex redistribution would be in Area 3, which is already redistributing RIP into OSPF. Because RIP is here to stay in Building 2 of Area 3, you must take care in the transition to EIGRP.

Building C1 would be an easy transition with redistribution on Router C. Router C2 would need to cut over to EIGRP and turn on redistribution immediately.

At this point, the only routers running OSPF are the core Routers A, B, and C. The choice now is whether to transition one router at a time or to cut over in one fell swoop. If the transition is done at night, cutting over everything is the safer method. Only three routers are involved, and the transition of the areas has now proven successful. There is a backup to restore everything to the status quo, in case problems are encountered.

Although this appears rather dangerous, it is much safer and cleaner than configuring further redistribution on the three core routers.

3. *To prevent routing loops, the company has decided to implement distribution filters. Plan the content, location, and implementation of the filters.*

The routers must prevent feedback of routing information during redistribution. This can be done to avoid routing loops, which can result in the death of your network.

The purpose of the distribute lists is to prevent the routing processes from telling each other what they already know. You can impose split horizons between the routing protocols manually.

Simply, when Router A is configured to redistribute the EIGRP routes from Router A1 into OSPF, OSPF will have a filter permitting only those routes to be redistributed. In this way, the routes that OSPF sends into EIGRP will not be propagated back into OSPF. Likewise, these filters will be set on Routers B and C.

NOTE This example is given to discuss the role of distribute lists and filters. In reality, there is little chance for a routing loop between OSPF and EIGRP, because both protocols mark redistributed routes as external, using administrative distance to prefer internal routes.

Scenario 19-5, Part B—Configuration

The next step in your job is to deploy the network designed in Scenario 19-5, Part A. Use the solutions for Part A of Scenario 19-5 to direct you in configuring the addressing and summarization at the ABRs. For Scenario 19-5, Part B, perform the following tasks:

1. Configure basic EIGRP for Routers A, B, and C based on the design in Scenario 19-5, Part A.
2. Configure summarization on Routers A, B, and C.
3. Configure redistribution between OSPF and EIGRP on Routers A, B, and C.
4. Configure distribution filters on Routers A, B, and C.

Solutions to Scenario 19-5, Part B—Configuration

Example 19-5, Example 19-6, and Example 19-7 show the configurations for Scenario 19-5, Part B, given the preceding criteria.

Example 19-5 Scenario 19-5 Router A Configuration

```
!Answer to Question 1: Configure basic EIGRP
router eigrp 200
network 131.99.0.0
passive-interface Ethernet0
!!! This prevents updates from being sent out of interface
!!! ETHERNET0, which is in the OSPF domain
redistribute ospf 100 metric 1000 100 255 1 1500
!!! Previous configuration
interface Ethernet0
ip address 131.99.0.49 255.255.255.240
! 4 bits allows 14 hosts
interface E1
ip address 131.99.136.129 255.255.255.240
! 4 bits allows 14 hosts
! Answer to Question 2: Configure summarization
ip summary-address EIGRP 200 131.99.192.0 255.255.240.0
ip summary-address EIGRP 200 131.99.224.0 255.255.240.0
! EIGRP summary addresses are interface specific,
! which allows different summarized addresses to be
! advertised out of different interfaces
interface E2
```

continues

Example 19-5 *Scenario 19-5 Router A Configuration (Continued)*

```

ip address 131.99.144.129 255.255.255.240
! 4 bits allows 14 hosts
!Answer to Question 2: Configure summarization
ip summary-address eigrp 200 131.99.192.0 255.255.240.0
ip summary-address eigrp 200 131.99.224.0.255.255.240.0
!!
router ospf 100
network 131.99.0.49 0.0.0.0 area 0
! The wildcard mask 0.0.0.0 places the interface ETHERNET0 into area 0
network 0.0.0.0 255.255.255.255 area 1
! The wildcard mask 255.255.255.255 places all other interfaces into area 1
area 1 range 131.99.128.0 255.255.240.0
! All addresses from area 1 will be summarized into this one network
area 1 stub
! Area 1 is defined as a stub network that will not receive
! network information outside the autonomous system
!!! New OSPF Configuration
!Answer to Question 3: Configure redistribution
redistribute EIGRP 200 metric 50 subnets
summary-address 131.99.128.0 255.255.240.0
passive-interface e1
passive-interface e2
!Answer to Question 4: Configure filters
distribute-list 101 in
access-list 101 permit ip 131.99.192.0 0.0.15.255 any
access-list 101 permit ip 131.99.224.0 0.0.15.255 any
access-list 101 permit ip 131.99.0.0 0.0.0.255 any
access-list 101 deny ip any any log
!!! Prevents OSPF networks from being looped back into the OSPF process

```

Example 19-6 *Scenario 19-5 Router B Configuration*

```

!Answer to Question 1: Basic EIGRP configuration
Router eigrp 200
network 131.99.0.0
passive interface Ethernet0
!!! This prevents updates from being sent out of interface ETHERNET0,
! which is in the OSPF domain
!Answer to Question 3: Configure redistribution (OSPF into EIGRP)
redistribute ospf 100 metric 1000 100 255 1 1500
!
interface Ethernet0
ip address 131.99.0.50 255.255.255.240
! 4 bits allows 14 hosts
interface e1
ip address 131.99.192.193 255.255.255.240
! 4 bits allows 14 hosts

```

Example 19-6 *Scenario 19-5 Router B Configuration (Continued)*

```

!Answer to Question 2: Configure summarization for interface E1
ip summary-address eigrp 200 131.99.128.0 255.255.240.0
ip summary-address eigrp 200 131.99.224.0.255.255.240.0
! EIGRP summary addresses are interface specific, which allows different
! summarized addresses to be advertised out of different interfaces
interface e2
ip address 131.99.216.193 255.255.255.240
! 4 bits allows 14 hosts
!Answer to Question 2: Configure summarization for interface E2
ip summary-address eigrp 200 131.99.128.0 255.255.240.0
ip summary-address eigrp 200 131.99.224.0.255.255.240.0
! EIGRP summary addresses are interface specific, which allows different
! summarized addresses to be advertised out of different interfaces
!!
router ospf 100
network 131.99.0.50 0.0.0.0 area 0
network 0.0.0.0 255.255.255.255 area 2
area 2 range 131.99.192.0 255.255.240.0
area 2 stub
!!! New OSPF Configuration
!Answer to Question 3: Configure redistribution (EIGRP into OSPF)
redistribute eigrp 200 metric 50 subnets
summary-address 131.99.192.0 255.255.240.0
passive-interface e1
passive-interface e2
!Answer to Question 4: Configure filters
distribute-list 101 in
access-list 101 permit ip 131.99.128.0 0.0.15.255 any
access-list 101 permit ip 131.99.224.0 0.0.15.255 any
access-list 101 permit ip 131.99.0.0 0.0.0.255 any
access-list 101 deny ip any any log
!
!!! Prevents OSPF networks from being looped back into the OSPF process

```

Example 19-7 *Scenario 19-5 Router C Configuration*

```

!Answer to Question 1: Basic EIGRP configuration
Router eigrp 200
network 131.99.0.0
passive interface Ethernet0
!!! This prevents updates from being sent out of interface ETHERNET0,
!!! which is in the OSPF domain
!Answer to Question 3: Configure redistribution (OSPF into EIGRP)
redistribute ospf 100 metric 1000 100 255 1 1500
!
interface Ethernet0
ip address 131.99.0.51 255.255.255.240

```

continues

Example 19-7 *Scenario 19-5 Router C Configuration (Continued)*

```

! 4 bits allows 14 hosts
interface e1
ip address 131.99.224.225 255.255.255.240
! 4 bits allows 14 hosts
!Answer to Question 2: Configure summarization for interface E1
ip summary-address eigrp 200 131.99.192.0 255.255.240.0
ip summary-address eigrp 200 131.99.128.0.255.255.240.0
! EIGRP summary addresses are interface specific, which allows different
! summarized addresses to be advertised out of different interfaces
interface e2
ip address 131.99.232.225 255.255.255.240
!7 bits allows 127 hosts
!Answer to Question 2: Configure summarization for interface E2
ip summary-address eigrp 200 131.99.192.0 255.255.240.0
ip summary-address eigrp 200 131.99.128.0.255.255.240.0
! EIGRP summary addresses are interface specific, which allows different
! summarized addresses to be advertised out of different interfaces)
!!

router ospf 100
network 131.99.0.51 0.0.0.0 area 0
network 0.0.0.0 255.255.255.255 area 3
area 3 range 131.99.224.0 255.255.240.0
!!! New OSPF Configuration
!Answer to Question 3: Configure redistribution (EIGRP into OSPF)
redistribute eigrp 200 metric 50 subnets
summary-address 131.99.224.0 255.255.240.0
passive interface e1
passive interface e2
!Answer to Question 4: Configure filters
distribute-list 101 in
access-list 101 permit ip 131.99.128.0 0.0.15.255 any
access-list 101 permit ip 131.99.192.0 0.0.15.255 any
access-list 101 permit ip 131.99.0.0 0.0.0.255 any
access-list 101 deny ip any any log
!!! Prevents OSPF networks from being looped back into the OSPF process

```

Scenario 19-5, Part C—Verification and Questions

Answer the questions following Example 19-8. Use Example 19-8 as a reference when the question refers directly to this scenario. Although not all of these questions are directly tied to the previous scenario, they all probe foundational knowledge required by the technology examined in this scenario.

NOTE In the network from which these commands were captured, several administrative settings not mentioned in the scenario were configured. For instance, the **enable** password was configured. Any **show running-config** commands in the examples in this chapter might have other unrelated configuration.

Example 19-8 *The show and debug Output Screens for Scenario 19-5, Router A*

```

routerA#
routerA#show ip protocols
Routing Protocol is "ospf 100"
  Sending updates every 0 seconds
  Invalid after 0 seconds, hold down 0, flushed after 0
  Outgoing update filter list for all interfaces is
  Incoming update filter list for all interfaces is 101
  Redistributing: ospf 100, eigrp 200
  Address Summarization:
  Routing for Networks:
    131.99.0.49/32
    0.0.0.0
  Passive Interface(s):
    Ethernet1
    Ethernet2
  Routing Information Sources:
    Gateway          Distance      Last Update
    131.99.232.225    110          00:07:08
    131.99.216.193    110          00:07:08
  Distance: (default is 110)

Routing Protocol is "eigrp 200"
  Outgoing update filter list for all interfaces is
  Incoming update filter list for all interfaces is
  Default networks flagged in outgoing updates
  Default networks accepted from incoming updates
  EIGRP metric weight K1=1, K2=0, K3=1, K4=0, K5=0
  EIGRP maximum hopcount 100
  EIGRP maximum metric variance 1
  Redistributing: ospf 100 (internal, external 1 & 2, nssa-external 1 & 2)

  Redistributing: eigrp 200
  Automatic network summarization is in effect
  Address Summarization:
    131.99.224.0/20 for Ethernet1, Ethernet2
    131.99.192.0/20 for Ethernet1, Ethernet2
  Routing for Networks:
    131.99.0.0
  Passive Interface(s):
    Ethernet0

```

continues

Example 19-8 *The show and debug Output Screens for Scenario 19-5, Router A (Continued)*

```

Routing Information Sources:
  Gateway         Distance      Last Update
Distance: internal 90 external 170

RouterA# show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       I - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, ia - IS-IS inter area
       * - candidate default, U - per-user static route, o - ODR
       - periodic downloaded static route

Gateway of last resort is not set

    131.99.0.0/16 is variably subnetted, 6 subnets, 4 masks
O IA   131.99.224.0/20 [110/11] via 131.99.0.51, 00:08:04, Ethernet0
O IA   131.99.192.0/20 [110/11] via 131.99.0.50, 00:08:04, Ethernet0
C       131.99.0.48/25 is directly connected, Ethernet0
C       131.99.136.128/28 is directly connected, Ethernet1

RouterA# show ip eigrp topology
IP-EIGRP Topology Table for Autonomous System(200)/ID(131.99.144.129)

Codes: - Passive, A - Active, U - Update, Q - Query, R - Reply,
       R - Reply status

P 131.99.224.0/20, 1 successors, FD is 2585600
     via Redistributed (2585600/0)
P 131.99.192.0/20, 1 successors, FD is 2585600
     via Redistributed (2585600/0)
P 131.99.0.48/28, 1 successors, FD is 281600
     via Connected, Ethernet0
     via Rconnected (281600/0)
P 131.99.136.128/25, 1 successors, FD is 128256
     via Connected, Ethernet1
     via Rconnected (128256/0)
P 131.99.144.128/25, 1 successors, FD is 128256
     via Connected, Ethernet2
     via Connected (128256/0)
P 131.99.216.193/32, 1 successors, FD is 2585600
     via Redistributed (2585600/0)

```


Answer the following questions:

1. Which command is used to see whether a neighbor adjacency has been created?
2. What is the meaning of the acronym SIA, and where would you see it?
3. What is the difference between the neighbor table and the topology table?
4. How are default routes identified and advertised in EIGRP?
5. What command shows the different IP routing protocols running on a system?
6. How would you detect that an EIGRP neighbor has become unavailable?
7. In redistributing an IP routing protocol, how would you prevent routes from being propagated back into the originating protocol?
8. Which routing protocol supports multiple Layer 3 protocols?
9. When redistributing OSPF routes into EIGRP, how would you state the metric to be used in the new routing protocol?
10. Where would you see the successors for a route?
11. How would you turn on EIGRP?
12. How would you configure router summarization in EIGRP?
13. What command shows the route summarization configured in EIGRP?
14. What is the purpose of the bandwidth percentage configuration in EIGRP?
15. Where would you see the administrative distance for a route?
16. Which **debug** command is used to identify that there is a problem in creating an adjacency in EIGRP?
17. What is the purpose of the **subnet** parameter in the EIGRP and OSPF redistribution command?
18. What routes can be redistributed into another routing protocol?
19. What is the difference between the **in** and the **out** parameters when set on a distribute list?
20. What should you consider in configuring EIGRP across an NBMA cloud?

Solutions to Scenario 19-5, Part C—Verification and Questions

The answers to the questions for Scenario 19-5, Part C are as follows:

1. *Which command is used to see whether a neighbor adjacency has been created?*

The command that shows whether an adjacency has been formed is **show ip eigrp topology** .

The codes, which are part of this command's output, show the adjacency. Codes tell the state of the topology table entry. Passive and active refer to the EIGRP state, with respect to this destination. Update, query, and reply refer to the type of packet that is being sent. The codes are as follows:

- P—passive: Indicates no EIGRP computations are being performed for this destination.
- A—active: Indicates EIGRP is actively trying to find an alternative path for this destination. If the neighbors have to be polled for a good neighbor, this consumes resources on the network.
- U—update: Indicates that an update packet was sent to this destination.
- Q—query: Indicates that a query packet was sent to this destination.
- R—reply: Indicates that a reply packet was sent to this destination.
- r—reply status: A flag that is set after the software has sent a query and is waiting for a reply.
- successors: Gives the number of successors. This number corresponds to the number of next hops in the IP routing table.
- FD—feasible distance: Used in the feasibility condition check. If the neighbor's reported distance (the metric after the slash) is less than the feasible distance, the feasibility condition is met and that path is a feasible successor. When the software determines it has a feasible successor, it fails over to that FS instead of resolving a path using queries.
- replies: Gives the number of replies that are still outstanding (that have not been received) with respect to this destination. This information appears only when the destination is in active state.
- state: Gives the exact EIGRP state of this destination. It can be the number 0, 1, 2, or 3. This information appears only when the destination is active.
- via: Gives the IP address of the peer that told the software about this destination. The topology table lists the destination route and the number of successors or paths to that route. The via field lists these successors, and the remaining entries on the list are feasible successors.

2. *What is the meaning of the acronym SIA, and where would you see it?*

The acronym SIA stands for Stuck in Active, which means that an EIGRP neighbor has not replied to a query that was sent out. If the neighbor does not reply within a limited time, it is presumed dead and is flushed from the tables. This is to prevent a route from being permanently active as an alternative path is sought from unresponsive neighbors.

An error message will be generated to the screen, but it is possible to identify the problem by looking at the log files or issuing commands. In the **show ip eigrp topology** command, any neighbors that show an R have not yet replied (the active timer shows how long the route has

been active) and might be Stuck in Active. It is advisable to run this command several times; you begin to see which neighbors are not responding to queries (or which interfaces seem to have many unanswered queries). You should also examine this neighbor to see if it is consistently waiting for replies from any of its neighbors. Repeat this process until you find the router that is consistently not answering queries.

The problems are often on the link to this neighbor, or with memory or CPU utilization with this neighbor.

It is often better to reduce the query range instead of increasing the SIA timer.

3. *What is the difference between the neighbor table and the topology table?*

The neighbor table holds information about EIGRP neighbors, whereas the topology table lists all the routes known to have feasible successors.

4. *How are default routes identified and advertised in EIGRP?*

Default routes are shown with the address 0.0.0.0 and are advertised as an external address. This affects the administrative distance.

5. *What command shows the different IP routing protocols running on a system?*

The command that shows all the different IP routing protocols running on a system is **show ip protocols**. This shows the protocols, redistribution, and many other details.

6. *How would you detect that an EIGRP neighbor has become unavailable?*

The command **show ip eigrp topology** would show that there had not been a reply to packets sent to the neighbor. The command **show ip eigrp neighbors** shows the neighbors, how long they have been in the table, and the last time they were heard from.

7. *In redistributing an IP routing protocol, how would you prevent routes from being propagated back into the originating protocol?*

This would be done by creating a distribute list that permits only those routes that did *not* originate from the protocol into which the updates are being redistributed.

8. *Which routing protocol supports multiple Layer 3 protocols?*

EIGRP supports IP, AppleTalk, and IPX.

9. *When redistributing OSPF routes into EIGRP, how would you state the metric to be used in the new routing protocol?*

There are two ways of stating the metric to be used by the redistributed networks. The first is to include the metric on the redistribution command. This will give the stated metric to all routes sent to EIGRP from the routing process OSPF 200:

```
Router(config-router)#redistribute ospf 200 metric bandwidth delay
reliability load mtu
```

The second command assigns the same metric to all routes distributed into EIGRP from any source:

```
Router(config-router)#default-metric bandwidth delay reliability load mtu
```

10. *Where would you see the successors for a route?*

The successors to a route are held in the routing table, because it is the current next hop that is being used to forward traffic to the remote destination. The command that shows these successors is as follows:

```
Router#show ip route
```

11. *How would you turn on EIGRP?*

The command to turn on EIGRP is as follows. This starts the routing process:

```
Router(config)#router eigrp autonomous-system-number
```

The following command identifies the interfaces that will receive, send, and advertise updates for that process:

```
Router(config-router)#network network-number
```

12. *How would you configure router summarization in EIGRP?*

Route summarization can be configured in EIGRP. This is achieved at the interface level and allows great flexibility in configuration. The command is as follows:

```
Router(config-if)#ip summary-address eigrp autonomous-system-number network area
```

13. *What command shows the route summarization configured in EIGRP?*

The command **show ip route** shows the summarization.

14. *What is the purpose of the bandwidth percentage configuration in EIGRP?*

The purpose of the bandwidth percentage configuration in EIGRP is to limit the amount of bandwidth that can be taken by EIGRP routing traffic. By default, this is limited to 50 percent of the link. In NBMA clouds—in particular, Frame Relay—it might be advisable to tune this parameter, depending on the Committed Information Rate (CIR) of the links.

15. *Where would you see the administrative distance for a route?*

You can see the administrative distance of a route in the routing table. It is the number in brackets placed after the destination network. The first number is the administrative distance; the second is the metric.

16. Which **debug** command is used to identify that there is a problem in creating an adjacency in EIGRP?

This command:

```
Router#debug ip eigrp neighbor autonomous-system-number address
```

adds a filter to this command

```
Router#debug ip eigrp packets
```

and displays only IP packets for the stated process and address.

17. What is the purpose of the **subnets** parameter in the OSPF redistribution command?

The **subnets** parameter in the OSPF redistribution commands allows the subnets of major networks that are not directly connected to be redistributed into the protocol OSPF. Without this, only major networks will be distributed.

18. What routes can be redistributed into another routing protocol?

Static routes, directly connected routes, or routes learned dynamically from another routing protocol can be redistributed.

19. What is the difference between the **in** and the **out** parameters when set on a distribute list?

The **in** parameter determines routes entering a routing protocol from another routing protocol or from another router. The **out** parameter defines the routes that can be redistributed into another routing protocol either to another routing process or from an interface. For example, in the following commands, routes are redistributed from RIP into OSPF. Any outbound update sourced from RIP that is in the network 10.0.0.0 will be redistributed into OSPF:

```
Router(config)#access-list 1 permit 10.0.0.0 0.255.255.255
Router(config)#router ospf 100
Router(config-router)#default-metric 30
Router(config-router)#redistribute rip
Router(config-router)#distribute-list 1 out rip
```

The **in** and **out** parameters can also be set against an interface or globally against all updates entering the router. However, the **out** parameter cannot be used at the interface level for link-state routing protocols, as the routing table is not sent in updates. For example,

```
Router(config)#access-list 1 permit 10.0.0.0 0.255.255.255
Router(config)#router ospf 100
Router(config-router)#distribute-list 1 in
```

20. What should you consider in configuring EIGRP across an NBMA cloud?

When configuring EIGRP over NBMA clouds, you should consider the nature of the link. Is it point-to-point or multipoint? Also consider overhead traffic associated with EIGRP and bandwidth utilization:

- Over a point-to-point interface, set the bandwidth to reflect the CIR of the permanent virtual circuit (PVC).
- Over multipoint Frame Relay, ATM, Switched Multimegabit Data Services (SMDS), and ISDN PRI, the bandwidth is divided equally among the links. The configuration should reflect the percentage of the available bandwidth.

If the PVCs have different CIRs, either convert the links to point-to-point or configure the bandwidth to be a multiple of the lowest available CIR by the number of PVCs.

It is important in point-to-multipoint configurations to turn off split horizons in order to ensure that the routing updates are propagated to the other links.

Scenario 19-6

Part A of Scenario 19-6 begins with some planning guidelines about the implementation of a new IS-IS network. This requires a sound addressing scheme and decisions about the location of the Level 2 routers to allow communication between the areas. BGP is the method of gaining access into the Internet, which requires some consideration of policy-based routing.

Part B of the scenario asks you to configure the three routers to implement the planned design and a few other features. Finally, Part C asks you to examine router command output to discover details about the current operation. Part C also lists some questions related to the user interface and protocol specifications.

Scenario 19-6, Part A—Planning

Your job is to deploy a new network with three sites, as shown in Figure 19-8.

This is a new network that is to run Integrated IS-IS with connectivity into the Internet, using BGP. In due time, the rest of the network will be transitioned into the IS-IS network. Use the addressing scheme presented Table 19-5. For Part A of this scenario, perform the following tasks:

1. Summarization is necessary between the areas. Give a brief explanation of where summarization would be implemented in the IS-IS network.
2. Using the addressing scheme provided, determine the type of IS-IS routing that should be implemented across the Frame Relay cloud.

3. Explain why the election of the DIS on the NBMA cloud would happen, and why it is important to ensure that Router A is the elected DIS.
4. Explain the method BGP will employ to connect to the ISP from Router A.
5. Explain how IS-IS on Router A will propagate a default route to the rest of the IS-IS domain that points to the ISP.

Figure 19-8 *Diagram for Scenario 19-6*

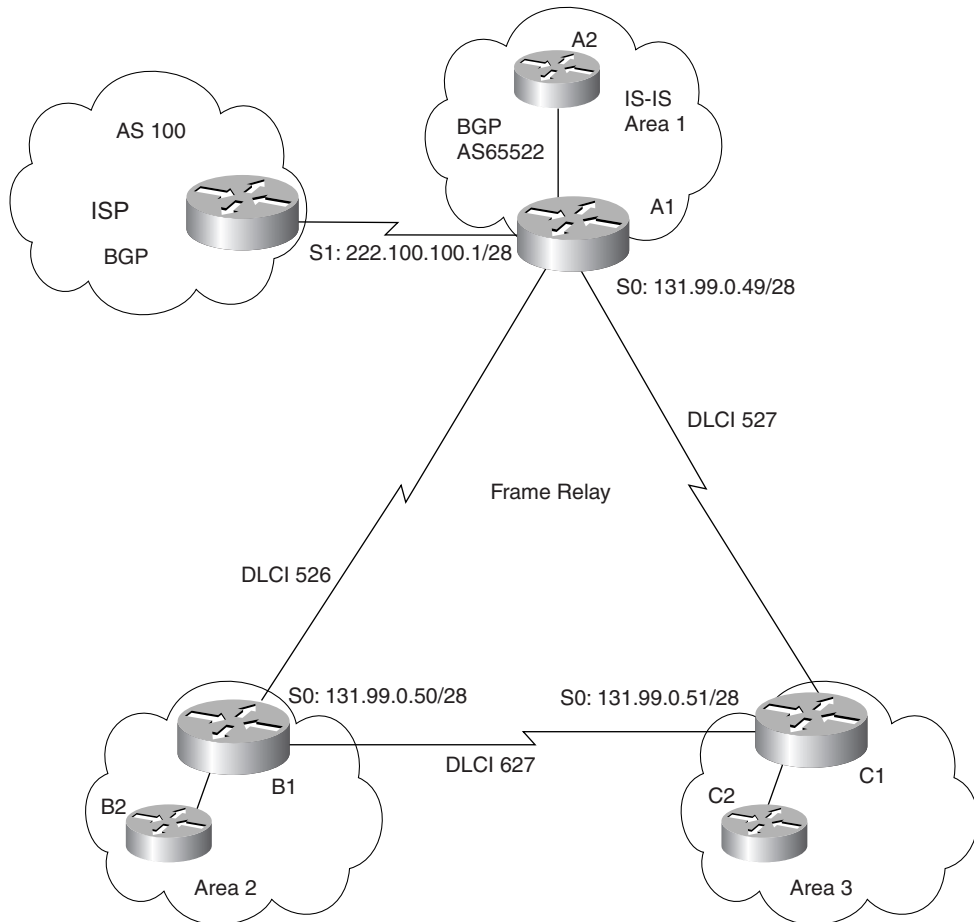


Table 19-5 Addressing Scheme for Scenario 19-6, Part A

Geographic Location of Subnet/Network	Subnet Number	Bit Pattern of Last Two Octets	Prefix	Subnet Mask
A				
Serial 0	131.99.0.48	00000000 0011 0000	/28	255.255.255.240
Serial 1	222.100.100.1	00000000 0010 0000	/28	255.255.255.240
Ethernet 0	131.99.136.128	10001000 1000 0000	/28	255.255.255.240
Ethernet 1	131.99.144.128	10010000 1000 0000	/28	255.255.255.240
B				
Serial 0	131.99.0.48	00000000 0011 0000	/28	255.255.255.240
Ethernet 0	131.99.192.192	11000000 1100 0000	/28	255.255.255.240
Ethernet 1	131.99.216.192	11011000 1100 0000	/28	255.255.255.240
C				
Serial 0	131.99.0.48	00000000 0011 0000	/28	255.255.255.240
Ethernet 0	131.99.224.224	11100000 1110 0000	/28	255.255.255.240
Ethernet 1	131.99.232.224	11101000 1110 0000	/28	255.255.255.240
Serial between A1 and A2	131.99.128.128	10000000 1000 0000	/26	255.255.255.192

Table 19-5 *Addressing Scheme for Scenario 19-6, Part A (Continued)*

Geographic Location of Subnet/Network	Subnet Number	Bit Pattern of Last Two Octets	Prefix	Subnet Mask
A1				
Ethernet 0	131.99.136.128	10001000 1000 0000	/28	255.255.255.240
Ethernet 1				
VLANs				
Floor 1	131.99.136.0	10001000 0 0000000	/25	255.255.255.128
Floor 2	131.99.137.0	10001001 0 0000000	/25	255.255.255.128
Floor 3	131.99.138.0	10001010 0 0000000	/25	255.255.255.128
Floor 4	131.99.139.0	10001011 0 0000000	/25	255.255.255.128
Floor 5	131.99.140.0	10001100 0 0000000	/25	255.255.255.128
Floor 6	131.99.141.0	10001101 0 0000000	/25	255.255.255.128
Floor 7	131.99.142.0	10001110 0 0000000	/25	255.255.255.128
A2				
Ethernet 0	131.99.144.128	10010000 1000 0000	/28	255.255.255.240
Ethernet 1				
VLANs				
Floor 1	131.99.144.0	10010000 0 0000000	/25	255.255.255.128
Floor 2	131.99.145.0	10010001 0 0000000	/25	255.255.255.128
Floor 3	131.99.146.0	10010010 0 0000000	/25	255.255.255.128
Floor 4	131.99.147.0	10010011 0 0000000	/25	255.255.255.128
Floor 5	131.99.148.0	10010100 0 0000000	/25	255.255.255.128
Floor 6	131.99.149.0	10010101 0 0000000	/25	255.255.255.128
Floor 7	131.99.150.0	10010110 0 0000000	/25	255.255.255.128

(continues)

Table 19-5 Addressing Scheme for Scenario 19-6, Part A (Continued)

Geographic Location of Subnet/Network	Subnet Number	Bit Pattern of Last Two Octets	Prefix	Subnet Mask
Serial between B1 and B2	131.99.192.128	11000000 10 000000	/26	255.255.255.192
B1				
Ethernet 0	131.99.192.192	11000000 1100 0000	/28	225.255.255.240
Ethernet 1				
VLANs				
Floor 1	131.99.200.0	11001000 0 0000000	/25	255.255.255.128
Floor 2	131.99.201.0	11001001 0 0000000	/25	255.255.255.128
Floor 3	131.99.202.0	11001010 0 0000000	/25	255.255.255.128
Floor 4	131.99.203.0	11001011 0 0000000	/25	255.255.255.128
Floor 5	131.99.204.0	11001100 0 0000000	/25	255.255.255.128
Floor 6	131.99.205.0	11001101 0 0000000	/25	255.255.255.128
Floor 7	131.99.206.0	11001110 0 0000000	/25	255.255.255.128
B2				
Ethernet 0	131.99.216.192	11011000 1100 0000	/28	225.255.255.240
Ethernet 1				
VLANs				
Floor 1	131.99.216.0	11011000 0 0000000	/25	255.255.255.128
Floor 2	131.99.217.0	11011001 0 0000000	/25	255.255.255.128
Floor 3	131.99.218.0	11011010 0 0000000	/25	255.255.255.128
Floor 4	131.99.219.0	11011011 0 0000000	/25	255.255.255.128
Floor 5	131.99.220.0	11011100 0 0000000	/25	255.255.255.128
Floor 6	131.99.221.0	11011101 0 0000000	/25	255.255.255.128
Floor 7	131.99.222.0	11011110 0 0000000	/25	255.255.255.128

Table 19-5 Addressing Scheme for Scenario 19-6, Part A (Continued)

Geographic Location of Subnet/Network	Subnet Number	Bit Pattern of Last Two Octets	Prefix	Subnet Mask
Serial between C1 and C2	131.99.224.128	11100000 10 000000	/26	255.255.255.192
C1				
Ethernet 0	131.99.224.224	11100000 1110 0000	/28	225.255.255.240
Ethernet 1				
VLANs				
Floor 1	131.99.224.0	11100000 0 0000000	/25	255.255.255.128
Floor 2	131.99.225.0	11100001 0 0000000	/25	255.255.255.128
Floor 3	131.99.226.0	11100010 0 0000000	/25	255.255.255.128
Floor 4	131.99.227.0	11100011 0 0000000	/25	255.255.255.128
Floor 5	131.99.228.0	11100100 0 0000000	/25	255.255.255.128
Floor 6	131.99.229.0	11100101 0 0000000	/25	255.255.255.128
Floor 7	131.99.230.0	11100110 0 0000000	/25	255.255.255.128
C2				
Ethernet 0	131.99.232.224	11101000 1110 0000	/28	255.255.255.240
Ethernet 1				
VLANs				
Floor 1	131.99.232.0	11101000 0 0000000	/25	255.255.255.128
Floor 2	131.99.233.0	11101001 0 0000000	/25	255.255.255.128
Floor 3	131.99.234.0	11101010 0 0000000	/25	255.255.255.128
Floor 4	131.99.235.0	11101011 0 0000000	/25	255.255.255.128
Floor 5	131.99.236.0	11101100 0 0000000	/25	255.255.255.128
Floor 6	131.99.237.0	11101101 0 0000000	/25	255.255.255.128
Floor 7	131.99.238.0	11101110 0 0000000	/25	255.255.255.128

Solutions to Scenario 19-6, Part A—Planning

Keeping the design as simple as possible, yet not making it so simple that the network cannot evolve, is a good practice. In these suggested answers, remember that many solutions are available. The reasoning behind this planning is to allow the optimum solution while maintaining the strength of the existing network. As in the Scenario 19-4 solution, the addressing scheme allows summarization at the core level of the network. The first few bits in the third octet indicate the summarization bits.

1. *Summarization is necessary between the areas. Give a brief explanation of where summarization would be implemented in the IS-IS network.*

Summarization in IS-IS is straightforward. The complexity is in the addressing scheme design. In Figure 19-8, the IP routes from Router A are summarized into Areas 2 and 3. Router A is capable of summarizing routes because the router is both a Level 1 and Level 2 router, acting as an ABR in OSPF.

Likewise, Router B can summarize routes 131.99.192.0 through to 131.99.207.254 with a mask summarized from /28 to /20 into Areas 1 and 3.

Router C can summarize routes 131.99.224.0 through to 131.99.239.254 with a mask summarized from /28 to /20 into Areas 1 and 2.

2. *Using the addressing scheme provided, determine the type of IS-IS routing that should be implemented across the Frame Relay cloud.*

Frame Relay is a WAN technology that has evolved beyond the point-to-point capabilities. When configured in a full mesh, the cloud is multiaccess, although each circuit is discrete and therefore not a true broadcast medium. However, IS-IS acknowledges only two types of network topologies, broadcast and point-to-point. If the network link is not a serial line connecting to a single router, IS-IS automatically defines the link to be broadcast.

3. *Explain why the election of the DIS on the NBMA cloud would happen and why it is important to ensure that Router A is the elected DIS.*

In a truly fully meshed configuration, it is important to consider which is to be elected as the DIS router. In this scenario, Router A must be chosen as the DIS because it is the only router that can communicate with the Internet. The DIS is determined by configuring the priority of the outgoing interface on the hub router.

An illustration of this configuration is shown in Example 19-9, which you see in the next section “Solutions to Scenario 19-6, Part B—Configuration.” The **frame-relay map ip** command maps the IP destination address to the outgoing DLCI and defines the interface as a broadcast interface. Integrated IS-IS uses the links as if they were truly broadcast links and elects a DIS.

The **frame-relay map clns** command maps to the CLNS process on the destination router. Without the second command, no routes appear in the IP routing table because CLNS does not receive the frames to populate the IP routing table. Remember that these are IP routes carried in the IS-IS routing protocol. It is IS-IS that updates the IP routing table.

4. *Explain the method that BGP will employ to connect to the ISP from Router A.*

Given that there is only one connection into the ISP, and that the IS-IS domain is a stub network from the viewpoint of BGP, use the simplest configuration: Set a static route from the ISP and a default route from the IS-IS domain. However, this company plans to eventually configure a dual-homed network to provide a redundant path from another router, and the administrators wish to ensure that the BGP configuration works. Therefore, you should opt for a BGP configuration on Router A. Because the company is using a private autonomous system number, this number will be stripped from the updates when they arrive at the ISP.

5. *Explain how IS-IS on Router A will propagate a default route to the rest of the IS-IS domain that points to the ISP.*

Configure a default route that is to be propagated throughout the entire network so that the rest of the IS-IS network can see the path into the Internet.

Scenario 19-6, Part B—Configuration

The next step in your job is to deploy the network designed in Scenario 19-6, Part A. Use the solutions for Part A of Scenario 19-6 to direct you in configuring the addressing and summarization at the ABRs. For Scenario 19-6, Part B, perform the following tasks:

1. Configure IS-IS on Routers A, B, and C based on the design in Scenario 19-6 Part A.
2. Configure IS-IS across the NBMA Frame Relay cloud using broadcast IS-IS.
3. Configure IS-IS summarization on Routers A, B, and C.
4. Configure Router A for BGP with its neighbor in the ISP.
5. Advertise the company's networks into the ISP using the **network** command.
6. Configure the default network for the rest of the IS-IS domain toward the ISP.

Solutions to Scenario 19-6, Part B—Configuration

Example 19-9, Example 19-10, and Example 19-11 show the configurations for Scenario 19-6, Part B, given the preceding criteria. The examples show the complete configurations for each router in turn. Within the configuration for each router, the solutions to each question are highlighted and identified by using comments in the configuration.

Example 19-9 Scenario 19-6 Router A Configuration

```
Working Example for Router A
!
interface Ethernet0
 ip address 131.99.136.129 255.255.255.240
!Answer to Question 1, Configure IS-IS
```

continues

Example 19-9 *Scenario 19-6 Router A Configuration (Continued)*

```

ip router IS-IS
!
interface Ethernet1

ip address 131.99.144.129 255.255.255.240
!Answer to Question 1, Configure IS-IS
ip router isis
IS-IS circuit-type level-1
!
interface Serial0
ip address 131.99.0.49 255.255.255.240
!Answer to Question 1, Configure IS-IS
ip router isis
IS-IS circuit-type level-1-2

!Answer to Question 2,Configuring for NBMA
encapsulation frame-relay

clockrate 56000
frame-relay map clns 526 broadcast
!Map the layer 2 DLCI address in Frame Relay to the clns process of Router C
frame-relay map clns 527 broadcast
!Map the layer 2 DLCI address in Frame Relay to the clns process of Router B
frame-relay map ip 131.99.0.50 526 broadcast
!Map the layer 2 DLCI address in Frame Relay to the Destination IP address of Router B
frame-relay map ip 131.99.0.51 527 broadcast
!Map the layer 2 DLCI address in Frame Relay to the Destination IP address of Router C
!
interface Serial1
Ip address 222.100.100.1 255.255.255.240
!Answer to Question 1, Configure IS-IS
IS-IS circuit-type level-2
!
router IS-IS

net 49.0001.0000.0000.000a.00

!Answer to Question 6, Configure default network
! Creating the default network to the ISP
ip default-network 222.100.100.0
ip classless
!
!Configuring for BGP and using the network command to send routes to the ISP

!Answer to Question 4, Configure BGP
Router bgp 65522
!Answer to Question 5, Configure BGP network command

```

Example 19-9 *Scenario 19-6 Router A Configuration (Continued)*

```

Network 131.99.0.0 255.255.0.0

Neighbor 222.100.100.2 remote-as 100

```

Example 19-10 *Scenario 19-6 Router B Configuration*

```

Router B
!
interface Ethernet0
 ip address 131.99.192.193 255.255.255.240
!Answer to Question 1, Configure IS-IS
 ip router isis
 IS-IS circuit-type level-1
!
interface Ethernet1

 ip address 131.99.216.193 255.255.255.240
!Answer to Question 1, Configure IS-IS
 ip router IS-IS
 IS-IS circuit-type level-1

!
interface Serial0
 ip address 131.99.0.50 255.255.255.240
!Answer to Question 1, Configure IS-IS
 ip router isis
 IS-IS circuit-type level-2
!Answer to Question 2,Configuring for NBMA
!Configure for NBMA
 Encapsulation frame-relay
 clockrate 56000
 frame-relay map clns 526 broadcast
!Map the layer 2 DLCI address in Frame Relay to the clns process of Router C
 frame-relay map clns 627 broadcast
!Map the layer 2 DLCI address in Frame Relay to the clns process of Router A
 frame-relay map ip 131.99.0.49 526 broadcast
!Map the layer 2 DLCI address in Frame Relay to the Destination IP address of Router A
 frame-relay map ip 131.99.0.51 627 broadcast
!Map the layer 2 DLCI address in Frame Relay to the Destination IP address of Router C
!
!Answer to Question 1, Configure IS-IS
router isis

net 49.0002.0000.0000.000b.00
!Answer to Question 3, Configure IS-IS summarization
!Summarization
Summary-address 131.99.192.0 255.255.240.0
 ip classless
!

```

Example 19-11 *Scenario 19-6 Router C Configuration*

```

Router C
!
interface Ethernet0
 ip address 131.99.224.224 255.255.255.240
!Answer to Question 1, Configure IS-IS
 ip router isis
!
interface Ethernet1

 ip address 131.99.232.224 255.255.255.240
!Answer to Question 1, Configure IS-IS
 ip router isis
IS-IS circuit-type level-1
!
interface Serial0
 ip address 131.99.0.51 255.255.255.240
!Answer to Question 1, Configure IS-IS
 ip router isis
IS-IS circuit-type level-2
!Answer to Question 2,Configuring for NBMA
! Configuring for NBMA
 encapsulation frame-relay
 clockrate 56000
 frame-relay map clns 527 broadcast
!Map the layer 2 DLCI address in Frame Relay to the clns process  of Router A
 frame-relay map clns 627 broadcast
!Map the layer 2 DLCI address in Frame Relay to the clns process  of Router B
 frame-relay map ip 131.99.0.50 627 broadcast
!Map the layer 2 DLCI address in Frame Relay to the Destination IP address of Router B
 frame-relay map ip 131.99.0.49 527 broadcast
!Map the layer 2 DLCI address in Frame Relay to the Destination IP address of Router A
!
! !Answer to Question 1, Configure IS-IS
router isis

 net 49.0003.0000.0000.000c.00
!Answer to Question 3, Configure IS-IS summarization
!Summarization
Summary-address 131.99.224.0 255.255.240.0
 ip classless
!

```

Scenario 19-6, Part C—Verification and Questions

Answer the questions following Example 19-12. Use Example 19-12 as a reference when the question refers directly to this scenario. Although not all of these questions are directly tied to the

previous scenario, they all probe foundational knowledge required by the technology examined in this scenario.

NOTE In the network from which these commands were captured, several administrative settings not mentioned in the scenario were configured. For instance, the **enable** password was configured. Any **show running-config** commands in the examples in this chapter might have other unrelated configuration.

Example 19-12 *The show and debug Output Screens for Scenario 19-6, Router A*

```
A#show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, * - candidate default
       U - per-user static route, o - ODR

Gateway of last resort is not set

*    222.100.100.0/28 is subnetted, 1 subnets
C*    222.100.100.0 is directly connected, Serial1
      131.99.0.0/16 is variably subnetted, 6 subnets, 2 masks
i L2   131.99.224.0/20 [115/20] via 131.99.0.51, Serial0
i L2   131.99.192.0/20 [115/20] via 131.99.0.50, Serial0
C      131.99.0.48/28 is directly connected, Serial0
C      131.99.136.128/28 is directly connected, Ethernet0
C      131.99.144.128/28 is directly connected, Loopback1
i L2   131.99.216.192/28 [115/20] via 131.99.0.50, Serial0
B      192.168.0.0/24 [20/0] via 222.100.100.2, 00:03:37

A#show isis database
IS-IS Level-1 Link State Database
LSPID                LSP Seq Num  LSP Checksum  LSP Holdtime  ATT/P/OL
0000.0000.000A.00-00* 0x00000029   0x18E9        1190          1/0/0
0000.0000.000A.01-00* 0x00000026   0x3178         545          0/0/0
0000.0000.000A.02-00* 0x00000003   0x705B         597          0/0/0
0000.0000.000A.03-00* 0x00000003   0x6961         791          0/0/0

IS-IS Level-2 Link State Database
LSPID                LSP Seq Num  LSP Checksum  LSP Holdtime  ATT/P/OL
0000.0000.000A.00-00* 0x00000029   0xD9BF         438          0/0/0
0000.0000.000B.00-00   0x00000009   0x4F25         529          0/0/0
0000.0000.000C.00-00   0x00000004   0xD198         581          0/0/0
0000.0000.000C.02-00   0x00000004   0x6EB2         529          0/0/0
```

continues

Example 19-12
The show and debug Output Screens for Scenario 19-6, Router A (Continued)

```

A#show clns neighbor

System Id      Interface  SNPA              State  Holdtime  Type Protocol
0000.0000.000B Se0      DLCI 526          Up     29        L2   IS-IS
0000.0000.000C Se0      DLCI 527          Up     9         L2   IS-IS


A#show ip bgp
BGP table version is 3, local router ID is 131.99.136.129
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: I - IGP, e - EGP, ? - incomplete


      Network          Next Hop           Metric LocPrf Weight Path
*> 131.99.0.0          0.0.0.0              0           32768 I
*> 192.168.0.0         222.100.100.2        0              0 100 I


A#show ip bgp neighbor
BGP neighbor is 222.100.100.2, remote AS 100, external link
Index 1, Offset 0, Mask 0x2
BGP version 4, remote router ID 192.168.0.1
BGP state = Established, table version = 3, up for 00:03:51
Last read 00:00:52, hold time is 180, keepalive interval is 60 seconds
Minimum time between advertisement runs is 30 seconds
Received 7 messages, 0 notifications, 0 in queue
Sent 7 messages, 0 notifications, 0 in queue
Prefix advertised 1, suppressed 0, withdrawn 0
Connections established 1; dropped 0
Last reset never
1 accepted prefixes consume 32 bytes
0 history paths consume 0 bytes
Connection state is ESTAB, I/O status: 1, unread input bytes: 0
Local host: 222.100.100.1, Local port: 11014
Foreign host: 222.100.100.2, Foreign port: 179


Enqueued packets for retransmit: 0, input: 0 mis-ordered: 0 (0 bytes)


Event Timers (current time is 0x1D6EF90):
Timer          Starts    Wakeups          Next
Retrans         8         0             0x0
TimeWait        0         0             0x0
AckHold         7         1             0x0
SendWnd         0         0             0x0
KeepAlive       0         0             0x0
GiveUp          0         0             0x0
PmtuAger        0         0             0x0
DeadWait        0         0             0x0


iss: 1569796080  snduna: 1569796256  sndnxt: 1569796256      sndwnd: 16209

```

Example 19-12 *The **show** and **debug** Output Screens for Scenario 19-6, Router A (Continued)*

```

irs: 1841527781 rcvnx: 1841527958 rcvwnd: 16208 delrcvwnd: 176

SRTT: 540 ms, RTT0: 3809 ms, RTV: 1364 ms, KRTT: 0 ms
MinRTT: 16 ms, maxRTT: 300 ms, ACK hold: 200 ms
Flags: higher precedence, nagle

Datagrams (max data segment is 1460 bytes):
Rcvd: 12 (out of order: 0), with data: 7, total data bytes: 176
Sent: 10 (retransmit: 0), with data: 7, total data bytes: 175

```

Answer the following questions:

1. Figure 19-8 shows three IS-IS areas. Which routers in this topology could be configured as Level 1 routers? State your reasons for your answers.
2. It is possible to configure the IS-IS level of routing at the interface level. Which routers in the topology could be configured in this manner? State the levels of routing that would be chosen, and give reasons for your answers.
3. What command is used to configure interface level routing?
4. What are the different methods of configuring IS-IS across an NBMA cloud? Which configuration does Cisco recommend, and why?
5. In examining the **show clns neighbor** command, identify the neighbors to Router A. When is the next Hello due from Router B, assuming that the interface is using the defaults?
6. The **show clns neighbor** command shows the Layer 2 address of the neighbors. Give the Layer 2 addresses for the neighbors and the name of the field that provided the information.
7. Which command reveals how long a neighbor has had an adjacency with the router and both the Layer 2 and Layer 3 addresses of the neighboring router?
8. Which command shows exactly the same information on every router in the same area that is operating at the same level of routing?
9. Which command reveals that the OL bit has been set, and what does this mean?
10. What is the purpose of the command **show isis spf-log** ?
11. Explain briefly the synchronization rule for BGP.
12. What does the command **show ip bgp** reveal about attributes?
13. What does the status code ***>** mean in the output of the **show ip bgp** command?
14. What is the purpose of a prefix list and when might it be used?

15. Why does Cisco advise that you do not redistribute dynamically learned routes from the IGP into the BGP process?
16. Many organizations want to connect to more than one ISP for redundancy. What are some of the concerns about following this strategy?
17. When and why would you issue the command **clear ip bgp *** ?
18. In the command **show ip bgp neighbors** , what does the notifications field show?

Solutions to Scenario 19-6, Part C—Verification and Questions

The answers to the questions for Scenario 19-6, Part C, are as follows:

1. *Figure 19-8 shows three IS-IS areas. Which routers in this topology could be configured as Level 1 routers? State your reasons for your answers.*

The routers within the areas, that is, A, A2, B, B2, C1, and C2, could be configured as Level 1 routers because they are internal routers to their area.

2. *It is possible to configure the IS-IS level of routing at the interface level. Which routers in the topology could be configured in this manner? State the levels of routing that would be chosen, and give reasons for your answers.*

The routers in the backbone would be configured as Level 1-2 routers. The serial interfaces forming a full mesh in the Frame Relay cloud have no communication with any Level 1 routers, so they could be configured as Level 2 at the interface. The Ethernet interfaces would need to be configured as Level 1 routers so that they could receive the Level 1 updates from the areas.

The routers A2, B2, and C2 could be configured as Level 1 routers.

3. *What command is used to configure interface level routing?*

The command used to configure the IS-IS router level on the interface is **isis circuit-type {level-1| level-1-2| level-2-only}** .

4. *What are the different methods of configuring IS-IS across an NBMA cloud? Which configuration does Cisco recommend, and why?*

IS-IS acknowledges only two types of network topologies, broadcast and point-to-point. If the network link is not a serial line connecting to a single router, IS-IS automatically defines the link to be broadcast.

WAN technology has evolved beyond point-to-point capabilities. When an NBMA cloud is configured in a full mesh, the cloud is multiaccess, although each circuit is discrete and therefore not a true broadcast medium.

In a true hub-and-spoke configuration, it is important that the hub router is elected as the DIS router, because it is the only router that can communicate with the other routers. This is done by configuring the priority of the outgoing interface on the hub router.

You are recommended to configure subinterfaces and a point-to-point IS-IS network. Avoid NBMA multipoint topologies. They can be made to work, but they require complex configuration and do not work with the strengths of the routing protocol.

5. *In examining the **show clns neighbor** command, identify the neighbors to Router A. When is the next Hello due from Router B, assuming that the interface is using the defaults?*

The neighbors shown in the **show clns neighbor** command are identified by the system ID. Because the command is issued at Router A, the neighbors shown are 0000.0000.000B and 0000.0000.000C.

6. *The **show clns neighbor** command shows the Layer 2 address of the neighbors. Give the Layer 2 addresses for the neighbors and the name of the field that provided the information.*

The field labeled SNPA shows the Layer 2 addresses of the neighbors. Because the links are using Frame Relay, the Layer 2 addresses are DLCI addresses. The neighbor addresses are DLCI 526 and DLCI 527.

7. *Which command reveals how long a neighbor has had an adjacency with the router and both the Layer 2 and Layer 3 addresses of the neighboring router?*

The **show clns neighbor detail** command shows the uptime as well as the subnetwork point of attachment (SNPA), which is the Layer 2 address, and the IP address.

8. *Which command shows exactly the same information on every router in the same area that is operating at the same level of routing?*

The **show isis database** command shows the same information on every router in the same area that is running the same level of routing. This is because the routing protocol is a link-state protocol that creates a database from every router's updates, as opposed to the distance vector method of sending a routing table to neighbors.

9. *Which command reveals that the OL bit has been set, and what does this mean?*

The **show isis database** command shows that the Overload (OL) bit has been set. The OL bit indicates that the router has an incomplete database because of memory overload and is therefore not used for transit data.

10. *What is the purpose of the **show isis spf-log** command?*

The **show isis spf-log** command shows how often and why the router has run a full shortest path first (SPF) calculation. Use the **show isis spf-log** command in EXEC mode.

11. *Explain briefly the synchronization rule for BGP.*

A simple rule states that before iBGP can propagate a route into another autonomous system it must hand the route over to eBGP. The route must be totally known within the autonomous system. That is, the Internal Gateway Protocol (IGP) or internal routing protocol must be synchronized with BGP.

This is to ensure that if traffic is sent into the autonomous system, the interior routing protocol can direct it to its destination. This rule is on by default and should be turned off only if all routers in the autonomous system are running BGP.

12. *What does the command **show ip bgp** reveal about the attribute?*

The **show ip bgp** command shows all the values of all the attributes and their status. Therefore, this is a good command to verify any configurations that change attributes to tune the system and to effectively manage the traffic flow to and from the autonomous system.

13. *What does the status code ***>** mean in the output of the **show ip bgp** command?*

The status code shows the status of the table entry. The status is displayed at the beginning of each line in the table. It can be one of the following values:

s—The table entry is suppressed.

*—The table entry is valid.

>—The table entry is the best entry to use for that network.

Therefore, ***>** indicates the entries are valid and the best path according to BGP.

14. *What is the purpose of a prefix list, and when might it be used?*

Prefix lists are used for filtering prefixes. From Cisco IOS software version 11.2, ISPs were given prefix lists, which are a more efficient form of filtering. Prefix lists filter based on the prefix of the address. This option was made a part of the general release IOS in version 12.0. Prefix lists are used on network numbers, not autonomous system paths, for which access lists are used.

15. *Why does Cisco advise that you do not redistribute dynamically learned routes from the IGP into the BGP process?*

Redistributing dynamically learned routes from the IGP is not advised because there is a great reliance on the IGP table. It is imperative that external routes carried in iBGP are filtered out; otherwise, routing loops are generated when BGP routes are fed into IGP, only to be advertised back into BGP farther down the network. Also, an IGP causes instability of BGP advertised routes. IGP will not handle the full Internet routing table.

16. *Many organizations want to connect to more than one ISP for redundancy. What are some of the concerns about following this strategy?*

The following are some concerns about connecting to more than one ISP:

- Each provider might not be propagating the same routes into or from the Internet. If the providers are sending subsets of the required routes, there could be a major problem with connectivity if the link to one of the providers fails.
- If you are connected to two different providers, your autonomous system could become a transit autonomous system between the ISPs. This could happen if a router in the autonomous system of one provider sees a path to a destination via the other provider's autonomous system, and your autonomous system gives the best route to the autonomous system of the other provider.

Configuration at the ISP level is the solution to these concerns and is dealt with when setting up the service. Therefore, it is important that you raise the need for multihoming during the negotiations with your ISP so that the ISP is aware of the need for additional configuration.

17. *When and why would you issue the command **clear ip bgp *** ?*

After configuration changes in BGP, it is necessary to reset the TCP session between neighbors. This can be forced with the following command:

```
Router(config-router)#clear ip bgp { * | address } [soft [in | out]]
```

This command disconnects the session between the neighbors and reestablishes it using the new configuration that has been entered. The soft option does not tear down the sessions, but resends the updates. The **in** and **out** options allow the configuration of inbound or outbound soft updates. The default is for both.

18. *In the command **show ip bgp neighbors** , what does the notifications field show?*

The notifications field shows the number of error messages that the router has sent to this peer.

Part VIII: Appendixes

Appendix A Answers to Chapter “Do I Know This Already?” Quizzes and Q&A Sections

Appendix B Suggested Reading



Answers to Chapter “Do I Know This Already?” Quizzes and Q&A Sections

Chapter 1

“Do I Know This Already?” Quiz

1. *a, c, d*
2. *a*
3. *b*
4. *d*
5. *b*
6. *a, c*
7. *a*
8. *d*
9. *a*
10. *d*
11. *a*
12. *a, d*

Q&A

1. *In the routing table, a field indicates the source of the routing information. If the field showed the letter C, what would this mean?*

A field showing the letter C would mean that the network is directly connected.

2. *In the routing table, how is the next hop indicated?*

In the routing table, the next hop is indicated by the word *via* followed by an IP address. This is the address of the next logical hop.

3. *Cisco distinguishes between the routing and the switching functions. What is the difference?*

The routing function is how the router learns the logical topology of the network. It decides whether the datagram can be routed, which path to select if there is a choice, and to which outgoing interface to queue the datagram. It operates at Layer 3 of the OSI stack.

The switching function is the forwarding of the frame from the inbound interface to an outbound interface. It operates at Layer 2 of the OSI stack, not at Layer 3 like a router, because the routing or Layer 3 function has already been completed.

4. *Name the interior IP routing protocols that send the mask with the routing update.*

The interior IP routing protocols that send the mask with the routing update are EIGRP, OSPF, IS-IS, and RIPv2.

5. *Does VLSM require a classful or classless routing protocol, and why?*

VLSM requires a classless routing protocol because it needs the subnet mask to be sent with the update.

6. *State one of the characteristics of a classful routing protocol.*

The characteristics of a classful routing protocol are as follows:

- It summarizes at the network boundary.
- Routes exchanged between foreign networks are summarized to the NIC number.
- Within the same network (IANA classful network), subnet routes are exchanged by routers.
- All the interfaces on all the routers within a NIC number must share the same subnet mask.
- VLSM is not possible within the network

7. *What is the command to show whether a specific network, such as 141.131.6.16, is present in the routing table?*

The command to show whether a specific network, such as 141.131.6.16, is present in the routing table is as follows:

show ip route 141.131.6.16

8. *State one major difference between a classful protocol and a classless routing protocol.*

Major differences between classful and classless routing protocols include:

- The capability to use VLSM
- The capability to summarize at an administratively defined boundary, as opposed to summarizing at the classful boundary
- The capability to maximize the logical address space

9. *Describe briefly the difference between a routing and routed protocol.*

The routed protocol is the Layer 3 protocol used to transfer data from one end device to another across the network. The routed protocol is the Layer 3 datagram that carries the application data in addition to the upper-layer information.

The routing protocol is the protocol used to send updates between the routers about the networks that exist in the organization, thereby allowing the routing process to determine the path of the datagram across the network.

10. *Describe the processes used to build and maintain the routing table.*

There are three steps involved in building and maintaining the routing table. These three processes are independent and include:

- The routing protocol, which actually sends the information about the routes or networks within the autonomous system, such as RIPv1, IGRP, EIGRP
- The routing table, which receives updates from the routing protocol and provides the forwarding process with information on request
- The forwarding process, which determines which path to select from the routing table in order to forward a datagram

11. *Give a brief explanation of the switching function.*

The switching function does the following:

- Checks the incoming frame for validity
- Checks whether the frame is addressed (at Layer 2) to the router
- Checks whether the frame is within the scope of the framing criteria (too big or too small)
- Checks whether the frame passes CRC
- Strips the Layer 2 header and trailer from the frame and checks the destination address against the cache entries
- Creates the appropriate frame header and trailer (if there is an entry in cache for the destination address) and forwards the frame to the outbound interface queue

12. *What information is contained in the fast switching cache?*

If fast switching is enabled, the datagram is then examined again, and an entry is put into a route cache. The entry in this cache consists of the following:

- An IP prefix
- The output interface
- The link-layer header to be used in forwarding the datagram

13. *When would you use the command **no auto-summary** ?*

Some routing protocols, such as BGP, RIPv2, and EIGRP, summarize at the network boundary automatically. Summarization within the NIC number boundary must be configured manually, and therefore autosummarization must first be disabled with the **no auto-summary** command.

14. *When does OSPF send updates, and do they contain the entire routing table?*

Each OSPF router refreshes its LSA every 30 minutes. An LSA will be sent earlier if there is an adjacency change.

15. *Network convergence is when a network domain has learned about the new network topology after a change has occurred in the domain. What is considered a network change that would require network convergence?*

The accuracy of the table will be affected by how quickly it responds to changes in the network. These changes include the following:

- Learning new networks
- Learning a better path to an existing network
- Learning that a network is no longer available
- Learning an alternative route to a network

16. *When would you consider using a static route in your network?*

A static route would be configured for the following reasons:

- Links that have very low bandwidth, such as dialup links
- The administrator needs control over the link
- The link is a backup to the dynamically learned route
- There is only one path to the remote network, such as a stub network

- The router has very limited resources and cannot run a routing protocol
- The administrator needs to control the routing table to allow a classful protocol and a classless routing protocol to populate the routing table

17. *When would you consider using a default route in your network?*

Occasions for using a default route include:

- Connecting to the autonomous system from a stub network
- Connecting to the Internet

18. *What command is used to configure ODR on a hub router?*

router odr is the only command needed to be configured on the hub router.

19. *When would you consider using a floating static route in your network?*

Floating static routes are used when a backup route is required.

20. *Describe some of the characteristics of classless routing protocols.*

Classless routing protocols include OSPF, EIGRP, RIPv2, IS-IS, and BGP.

- The characteristics of a classless routing protocol are:
- Router interfaces within the same network can have different subnet masks (VLSM).
- They support the use of classless interdomain routing (CIDR).
- Some routes can be summarized within the major NIC number. This is done manually.

Chapter 2

“Do I Know This Already?” Quiz

1. *a*
2. *b*
3. *d*
4. *c*
5. *a*
6. *b*

- 7. *a*
- 8. *b*
- 9. *d*
- 10. *c*
- 11. *a, c, d*
- 12. *d*

Q&A

- 1. *Identify one criterion to help determine a subnet mask for classless addressing when designing a network-addressing scheme.*

Questions to ask include the following:

- How many networks are there in the network?
 - How many hosts are there on the largest subnet?
- 2. *With a classless address of 204.1.64.0/20, what is the range of classful addresses that are included in the address? Write your answer in dotted decimal and the third octet in binary notation.*

The address 204.1.64.0 /20 includes the Class C addresses 204.1.64.0 to 204.1.79.0; this is illustrated in both dotted decimal and binary notation in the following table.

Binary Notation	Decimal Notation
01000000	204.1.64.0
01000001	204.1.65.0
01000010	204.1.66.0
01000011	204.1.67.0
01000100	204.1.68.0
01000101	204.1.69.0
01000110	204.1.70.0
01000111	204.1.71.0
01001000	204.1.72.0
01001001	204.1.73.0
01001010	204.1.74.0

Binary Notation	Decimal Notation
01001011	204.1.75.0
01001100	204.1.76.0
01001101	204.1.77.0
01001110	204.1.78.0
01001111	204.1.79.0

3. *What is a discontinuous network?*

A discontinuous network is a network in which a classful network is separated by another classful network. Therefore, the original classful network is no longer contiguous because an intervening Internet number has segmented or divided it.

4. *For VLSM to be available as a design option in the network, what characteristic must the routing protocol possess?*

The routing protocol must send the prefix or subnet mask as part of the routing update.

5. *If summarization is to be implemented in the network, name one design criterion for the addressing scheme that must be in place.*

For VLSM to work, the addressing scheme must be hierarchical, allowing the upstream devices to share the same high-order bits as the downstream devices.

6. *If the host portion of a subnet has been used to identify end devices, can that subnet be used again for VLSM?*

It is not possible to use a subnet for addressing hosts or to further subnet the network using VLSM. The addresses would be seen as duplicate addresses.

7. *Give one example of when route summarization would not be a good solution.*

Route summarization is not useful in the following circumstances:

- There are discontinuous networks in the organization.
- A specific subnet needs to be seen throughout the network.
- The addressing scheme does not support summarization. No common high-order bits are shared in the network-addressing scheme.
- Access lists require detailed information, which summarization suppresses.

8. *Give one reason for implementing route summarization.*

Route summarization is useful for the following reasons:

- To keep the routing tables small
- To keep the network overhead low
- To hide the network details from the rest of the organization
- To prevent flapping links from affecting the rest of the network

9. *Given an address of 133.44.0.0 and a prefix mask of /25, how many networks can be addressed, and how many hosts can exist on each network? Write the first and last possible subnets in binary and decimal notation.*

For the network address of 133.44.0.0, the subnet mask of 255.255.255.128 would enable you to address 510 subnets with 126 hosts on each subnet. This complies with the subnetting rule of not allocating addresses with all 0s or all 1s. The following table illustrates the first and last subnet in their binary and decimal notation formats.

Binary Notation	Decimal Notation
00000000.10000000	133.44.0.128
11111111.00000000	133.44.255.0

10. *What class of address is 131.188.0.0, and how many hosts can be addressed if no subnetting is used?*

131.188.0.0 is a Class B address and can address more than 65,000 hosts on one network if no subnetting is used.

11. *Write out the decimal notation of the following subnet mask presented in the binary notation of 11111111.11111111.11111111.11110000.*

The decimal notation of the subnet mask would be 255.255.255.248.

12. *Is 201.111.16.0/20 a valid subnet mask?*

Yes, this is a valid mask, and it will provide 16 consecutive Class C addresses to the organization.

13. *Briefly define route summarization.*

Route summarization is the method of including many subnets in a few routing entries.

14. *What sort of design scheme does route summarization require?*

Route summarization requires a hierarchical addressing scheme.

15. *In route summarization, in which direction is the network/host boundary in the subnet mask moved?*

In route summarization, the subnet mask is shifted to the left.

16. *Explain how summarization allows for smaller routing tables.*

Summarizing is the consolidation of multiple routes into a single advertisement.

17. *What is the subnet mask for a /21 prefix?*

The subnet mask is 255.255.248.0.

18. *What is the default subnet mask for the IP address 192.18.16.15?*

The subnet mask is 255.255.255.0.

19. *State whether 131.104.0.0/13 is an example of CIDR or VLSM routing.*

The address 131.104.0.0/13 is an example of CIDR routing because it is summarizing Class B addresses within the Internet routing tables.

20. *State how many classful addresses are summarized in the address 131.104.0.0/13.*

Eight Class B addresses are summarized in the address 131.104.0.0/13. The range of addresses is 131.104.0.0–131.111.0.0/16.

Chapter 3

“Do I Know This Already?” Quiz

1. *a, b, d*
2. *a*
3. *c*
4. *b*
5. *b, c*
6. *b, d*
7. *d*

- 8. *b*
- 9. *a*
- 10. *c*
- 11. *a*
- 12. *d*
- 13. *d*
- 14. *c*
- 15. *b*

Q&A

1. *What is an Inside Global address, and when is it used?*

The Inside Global address connects your organization indirectly to the Internet. Typically, these are the addresses provided by the ISP. These addresses are propagated outside the organization. They are globally unique and are the addresses used by the outside world to connect to inside the organization. Simply explained, they are the addresses that define how the *inside* addresses are seen *globally* by the outside.

2. *When is the Inside Local address used?*

The Inside Local address is the address that allows every end device in the organization to communicate. Although these addresses are unique within the organization, they are probably not globally unique. They may well be private addresses that conform to RFC 1918. They are the *inside* addresses as seen *locally* within the organization.

3. *Explain the difference between the Outside Global address and the Outside Local address.*

The Outside Global address is the Internet address (all the addresses outside the domain of the organization). They are the *outside* addresses as they appear to the *global* Internet.

The Outside Local address, however, is external to the organization. This is the destination address used by a host inside the organization connecting to the outside world. This will be the destination address of the packet propagated by the internal host. This is how the *outside* world is seen *locally* from inside the organization.

4. *Give one example of when NAT might be employed.*

NAT is useful in the following circumstances:

- To connect organizations that use address space issued to other organizations to the Internet
- To connect organizations that use private address space defined in RFC 1918 and want to connect to the Internet
- To connect together two organizations that have used the same private address, in line with RFC 1918
- When the organization wants to hide its addresses and is using NAT as part of firewall capabilities or is using additional security features

5. *Explain what PAT is.*

Port Address Translation (PAT) translates different local addresses (within the organization) into one address that is globally significant for use on the Internet. The additional identifier of a TCP or UDP port unravels the multiple addresses that have been mapped to single addresses. The uniqueness of the different local addresses is ensured by the use of the port number mapped to the single address.

6. *Why is NAT often configured on the organization's firewall?*

NAT performs Network Address Translation from private addressing to global addressing. This is required only when connecting to the outside world when a globally unique address is required. A firewall is installed on the edge of the autonomous system to protect the organization from unauthorized access. Therefore, both applications would reside on the edge of the network as it connects into the Internet.

7. *Give one example of when private addressing would be a good solution for an organization.*

Private addressing was designed as an addressing method for an organization that has no intention of ever connecting to the Internet. If Internet connectivity were not required, there would be no requirement for a globally unique address from the Internet. The individual organization could address its network without any reference to the Internet, using one of the address ranges provided. Private addressing is often used now by companies that do not want to make a submission for an address from the IANA or do not want the labor of stretching the small allocation of addresses that they might receive. Instead, they can use a Class A address for their organization and use NAT to connect to the outside world.

8. *Why does summarization need a hierarchical addressing structure?*

Summarization of IP addresses needs a hierarchical addressing structure to hide the downstream subnets. A summarized address is where many subnets that share the same leftmost bits of the IP address might be represented by a single subnet. This can only occur when the minor subnets are physically downstream of the summarized network.

9. *When might you use Destination Address Rotary Translation?*

You use destination address rotary translation for traffic entering the organization from the outside. The destination address is matched against an access list, and the destination address is replaced by an address from the rotary pool. This is used only for TCP traffic, unless other translations are in effect. This is required when creating TCP connections that require a connection into the network. It allows Telnet, FTP, and other connection-oriented sessions.

10. *Why do NAT and private addressing tend to be implemented together?*

NAT performs Network Address Translation and is implemented when private addresses need to be translated into a global address to allow connectivity to the Internet with a globally unique address.

11. *Which routing protocols for IPv6 does Cisco IOS support?*

Cisco supports RIPng, BGP-4+, IS-IS, and OSPF in Cisco IOS 12.2T.

12. *How many bits are there in the IPv6 address space?*

There are 128 bits in the IPv6 address space.

13. *Explain why it is important to remember to identify how many hosts and subnets are required when designing an IPv4 network.*

It is important to determine how many hosts and subnets are required so that you can adequately address the network, allowing for the physical topology of the network to marry the logical topology and to build in a hierarchical addressing scheme that can be summarized to add efficiency into the network.

14. *Why does IPv6 addressing allow for more effective security and QoS to be implemented?*

IPv6 addressing allows for more effective security and QoS to be implemented because the address space allows for end-to-end connectivity. Without having to traverse firewalls and NAT servers, both security and QoS are much easier to create and maintain.

15. *Give one reason it might be advisable to implement private addressing.*

The reasons for addressing your organization’s network using private addressing include the following:

- There is a shortage of addressing within the organization.
- You need security. Because the network must go through a translation gateway, it will not be visible to the outside world.

- There is an ISP change. If the network is connecting to the Internet through an ISP, the addresses allocated are just on loan or are leased to your organization. If the organization decides to change its ISP, the entire network will have to be readdressed. If the addresses provided define just the external connectivity and not the internal subnets, however, readdressing is limited and highly simplified.

16. *In the following address*

4021:0000:240E:0000:0000:0AC0:3428:121C

which part of the address is autoconfigured?

The MAC address 0AC0:3428:121C provides the system ID for the address.

17. *Why is IPv6 multicasting more efficient than IPv4 broadcasting?*

IPv4 broadcasting forces every end system to interrupt its process to view the incoming packet, because the broadcast address is an address to every system. This saturates not only the media resources, but also the resources of the end systems. Multicasting is an address to multiple end systems, but only to those systems that need to receive the packet.

18. *State one of the main benefits of IPv6.*

The main benefits of IPv6 are as follows:

- Larger address space, allowing for a larger number of systems that can be globally addressed and a more scalable network
- Increased address space, allowing for a deeper hierarchical structure
- Simplified header, allowing for greater routing efficiency and, thus, network performance
- Policies for network architecture flexibility, allowing evolution and growth of the protocol
- Support for routing and route aggregation
- Simple administration through serverless autoconfiguration, the ability to renumber with ease, multihoming—all of which allow a level of plug-and-play support
- Security using IP Security (IPSec) support for all IPv6 devices
- Support for Mobile IP and mobile computing devices (direct-path)
- Multicast support built into the protocol using a greater number of addresses and efficient mechanisms

19. *What is an IPv6 extension header?*

The IPv6 extension header is the same as the Options field in the IPv4 header. However, instead of including the Options field within the header as IPv4 does, IPv6 attaches the Options field to the end of the header, indicating with the Next Header field whether there is something to additional to process. This speeds up the processing and also allows for protocol evolution, because many extension fields can be chained together.

20. *What are the two most common methods of transitioning an IPv4 network to an IPv6 network?*

Dual stacks and 6to4 tunneling are the two most common methods of transitioning an IPv4 network to an IPv6 network.

Chapter 4

“Do I Know This Already?” Quiz

1. *a, b, c*
2. *b*
3. *c*
4. *c*
5. *c, d*
6. *b*
7. *b*
8. *c*
9. *a*
10. *c*
11. *c*
12. *a, b, c, d*
13. *a*
14. *d*
15. *b*
16. *b, c, d*
17. *a*
18. *b*

Q&A

1. *Name one routing protocol that sends periodic updates.*

RIPv1, RIPv2, and IGRP all send periodic updates.

2. *What is an incremental update, and how often is it sent out?*

An incremental update is an update that is sent out only when there is a change in the network. It contains only the information about the change. The change could be either the loss of a network or the addition of a network. EIGRP sends out incremental updates, as does OSPF, IS-IS, and BGP.

3. *Distance vector routing protocols naturally summarize at which boundary?*

Distance vector routing protocols naturally summarize at the IANA or major network boundary. They do this by following the first octet rule.

4. *What is the algorithm used by distance vector protocols?*

Distance vector routing protocols use the Bellman Ford algorithm.

5. *Give three reasons why RIPv1 has problems working in a large network.*

RIPv1 has problems working in a large network because of the following reasons:

- It has a maximum hop count of 15.
- It sends updates of its routing table out of every interface every 30 seconds, which increases the network overhead and leads to link congestion.
- To avoid routing loops, it uses holddown and poison reverse, and thereby increases the time that it takes to propagate the changes in the network.

6. *What is the destination address of the distance vector periodic update in RIPv1?*

The destination address of the distance vector periodic update is 255.255.255.255 (the broadcast address).

7. *State two ways that a route is selected as the preferred path.*

A route is selected for the following reasons:

- Because it is the only available path
- Because the administrative distance is lower
- Because the metric is lower

8. *What is administrative distance?*

Administrative distance is the mechanism used by the routing process to select a path offered by two different routing protocols. The administrative distance is a set of values, in which a value is given to each IP routing protocol. This allows a hierarchy to be established so that when multiple protocols offer a path to the same remote network, one path can be chosen. The path that is chosen will be the one offered by the routing protocol with the lowest administrative distance. The administrative distance can be manually configured.

9. *If IGRP has three paths to a remote network in which each path has an equal metric, what will happen?*

If IGRP sees equal-cost paths to a remote network, it will load balance between those paths by default.

10. *A distance vector routing protocol uses the mechanism of poison reverse. What is poison reverse?*

When the routing process suspects that a route in its routing table is no longer valid, it sets the metric so high for that route that it renders it unusable. This metric will be propagated in the routing updates to other routers. This mechanism is used in triggered updates.

It is also used with split horizon. Instead of suppressing routes from being advertised out of the interface from which they were heard, it advertises the routes but sets the metric so high as to be rendered unusable. This is called split horizon with poison reverse.

11. *Name two distance vector routing protocols.*

Distance vector routing protocols include RIPv1, RIPv2, IGRP, and EIGRP (an advanced distance vector routing protocol).

12. *Describe the mechanism of split horizon.*

Split horizon is a routing technique in which information about routes is prevented from exiting the router interface through which that information was received. Split-horizon updates are useful in preventing routing loops.

13. *What is meant by the phrase routing by rumor?*

Distance vector routing protocols are said to be “routing by rumor” because information is never sent or flooded to all routers but rather sent to neighbors who in turn forward their routing table.

Distance vector protocols send periodic updates. These updates are sent to directly connected neighbors. The update is periodic because it waits for a timer to expire before it sends an update.

After receiving a neighbor's routing table, the router updates its own table and sends the modified table in subsequent updates.

14. *Why does the use of multicast addressing in RIPv2 overcome some of the limitations of RIPv1?*

The use of the multicast addressing saves network resources because all the nodes in the network can discard the packet at either Layer 2 or Layer 3 instead of taking the packet all the way up to the transport layer, where the port number is rejected. Although this saves system resources, the same amount of bandwidth is consumed on each link that the multicast traverses.

15. *Explain the use of holddown in distance vector routing protocols to create stability in the network.*

The use of holddown is crucial to network stability. The term *holddown* refers to when and how the routing process decides that a route in the routing table is no longer valid. After deciding that a network in the routing table is no longer available, the routing process waits for three routing updates (by default) before it believes a routing update with a less-favorable metric. Again, this is to prevent routing loops from generating false information throughout the network.

16. *What is the maximum hop count in RIPv1 and RIPv2?*

The maximum hop count is 15 with infinity stated at 16.

17. *Both EIGRP and IGRP use a composite metric. What are the main components of this metric?*

Bandwidth, delay, reliability, and load are the interface values that are used for IGRP and EIGRP.

18. *Explain briefly how RIPv2 differs from RIPv1.*

RIPv2 differs from RIPv1 in the following ways:

- The metric is a composite calculated from bandwidth, delay, loading, reliability, and MTU. In fact, although MTU was originally designed as part of the metric, it is tracked but not used in the calculation. It is possible to configure the use of all the calculated elements of the metric. If these are not configured, the system will use only bandwidth and delay by default.
- The hop count is 100 and is configurable to 255 (although this is not used as a metric, but to age out datagrams).
- The update timer is set by default to 90 seconds (three times that of RIPv1).
- Unequal-cost load balancing occurs on multiple paths.

- A more efficient packet structure is used.
- Autonomous systems are used to allow multiple processes within a routing domain, which allows the network to scale.

19. *What is meant by the term convergence?*

Convergence occurs when all the routers in the routing domain agree on the available routes. Convergence time is the time that it takes for every router’s routing table to synchronize after there has been a change in the network topology.

It is important to ensure that the time taken is as short as possible, because, while the routers disagree on the available networks, they cannot route data correctly or efficiently.

20. *Give the configuration commands to turn on the process for RIPv2.*

```
RouterA(config)#router rip  
RouterA(config-router)#version 2
```

Chapter 5

“Do I Know This Already?” Quiz

1. *b*
2. *b*
3. *a, b, c*
4. *a, d*
5. *a, c*
6. *c, d*
7. *b*
8. *a*
9. *d*
10. *a, c*
11. *d*
12. *d*
13. *b*
14. *a, d*
15. *d*

Q&A

1. *What is the routing algorithm used in OSPF?*

The Dijkstra algorithm is the routing algorithm used in OSPF.

2. *State one method by which a link-state routing protocol attempts to reduce the network overhead.*

Link-state routing protocols are used in larger networks because the method that they use to update the routing tables requires fewer network resources.

Link-state routing protocols attempt to reduce network overhead by:

- Using multicast addressing
- Sending triggered updates
- Sending network summaries infrequently if at all
- Using small packets from every router to describe their local connectivity, instead of the entire routing table

3. *What is the purpose of the Dijkstra algorithm?*

This is a routing algorithm that iterates on the length of path to determine a shortest path first tree (SPF tree). It is commonly used in link-state routing protocols to determine which route to use.

4. *Name two link-state IP routing protocols.*

OSPF and IS-IS are two link-state IP routing protocols.

5. *Name the TCP port used by BGP.*

TCP port 179 is used by BGP.

6. *State the metric used by OSPF.*

The metric cost can be manually configured to represent anything you want, though Cisco has programmed its systems to use the inverse of bandwidth as the default.

7. *How often does Integrated IS-IS send out new LSAs?*

Integrated IS-IS sends out LSAs every 15 minutes or whenever a change in the network is seen.

8. *State one way that OSPF is an improvement over RIPv1.*

OSPF is an improvement over RIPv1 for large networks because of the following:

- It uses bandwidth more efficiently, sending incremental updates.
- The updates are not broadcast as in RIPv1 but are directed to multicast addresses 224.0.0.5 and 224.0.0.6.
- It propagates changes in the network more quickly, with incremental updates and neighbor relationships.
- It is not limited in size by a maximum hop count of 15.
- It allows for variation in network size throughout the organization, using VLSM.
- It is capable of using the MD5 specification.
- The metric can be defined manually, allowing for greater sophistication in the path determination.
- It is more responsive to network changes, is flexible in network addressing and design, and scales to a larger size.

9. *State one key attribute of OSPF.*

The key attributes of OSPF include the following:

- Maintaining a connection-oriented relationship with other routers on the same physical segment. These are known as adjacent neighbors.
- Sending the minimum amount of information in an incremental update when there has been a change in the network. This allows for fast network convergence.
- Adding another level of hierarchy to the IP address by designing networks into areas.
- Using VLSM and summarization.
- Assigning specific functionality to different routers to streamline the process of communication change in the network.
- Operating within an organization as an interior routing protocol.

10. *State one key attribute of IS-IS.*

The following are attributes of IS-IS:

- Routes CLNP traffic.
- Routes IP traffic.

- Allows VLSM and summarization.
- Uses network design of areas to limit CPU-intensive computation.
- Assigns functionality to routers to streamline the communication of network change. Level 1 routers deal with interarea updates, whereas Level 2 routers communicate between areas.
- Sends incremental updates to conserve both bandwidth and CPU.
- Operates within an autonomous system as an internal routing protocol.

11. *State one key attribute of BGP.*

The following are key attributes of BGP:

- Full routing updates are sent at the beginning of the session.
- Trigger or incremental updates are sent only after the initial setup.
- Connections between BGP routers are maintained by periodic hellos. The Hello protocol is connection-oriented, using TCP, port 179.
- It uses the hierarchical structure of autonomous systems.
- It has a complex metric, called attributes, by which traffic paths can be manipulated.

12. *What is the default hello update timer for IS-IS on broadcast media?*

The update hello timer for IS-IS on broadcast media is every 10 seconds.

13. *On a broadcast link, how long does OSPF wait by default before it determines that a neighbor is dead?*

On a broadcast link, OSPF would wait 40 seconds before declaring the neighbor as unavailable or dead. This would result in the creation and flooding of an LSA.

14. *What is iBGP?*

The acronym *iBGP* stands for internal BGP; it is used to traverse an autonomous system. An internal routing protocol is required to carry updates from one part of the autonomous system to another.

15. *When does OSPF send updates?*

Incrementally. The update only contains the network change.

However, 30 minutes after the last update is received, a compressed version of the table is propagated.

16. *When does BGP send updates?*

BGP sends incremental updates that contain only the network change.

17. *What is a topological database?*

Used by EIGRP and OSPF, this table records all the routes in the network before determining which will be entered into the routing table.

18. *What is an adjacent neighbor?*

A neighbor is a router that is directly connected to another router. They must also have the same mask and hello parameters. Once a neighbor relationship is formed, routing updates can be exchanged. An adjacent router is a router that has exchanged routing information with its neighbor, becoming an adjacent neighbor.

19. *What is a triggered update?*

A triggered update is when a routing update is sent asynchronously in response to a change in the network topology. If there is a change in the metric, the update is sent immediately without waiting for the update timer to expire.

20. *What is required for iBGP to operate?*

The iBGP is used to send routing information internally across an autonomous system, using it as a transit area to another autonomous system. The iBGP needs a fully meshed BGP network, but these routers do not need to be directly connected. The way the BGP updates can be sent to the other BGP routers, or the BGP data traffic can find the remote destination, is by listening to the interior IP routing protocol. Although the remote peer does not have to be directly connected, there is an entry in the routing table for the remote peer so that the routers can communicate with each other.

Chapter 6

“Do I Know This Already?” Quiz

1. *d*
2. *b*
3. *c*
4. *d*
5. *c*

- 6. *b*
- 7. *b*
- 8. *a*
- 9. *a*
- 10. *b*
- 11. *c*
- 12. *d*

Q&A

1. *What information is held in the topology table?*

The topology table holds a map of every link in the area. Every topology table in the area is the same. This is sometimes referred to as the link-state database.

2. *What command is used to determine manually which router on a LAN will become the DR?*

The *ip ospf priority number* command is used to determine the DR manually. The higher the priority, the greater the likelihood is of success.

3. *How many subnets are required in an OSPF configuration over a point-to-point network that has multiple connections?*

It is necessary to have one subnet per connection. Thus, if there are four point-to-point links, four subnets are required.

4. *State the different types of packets used to build a routing table for the first time.*

Five packets are used to build the routing table for the first time:

- **The hello packet** — This is used to find neighbors and to determine the designated and BDR. The continued propagation of the hello packet maintains the transmitting router in the topology database of those that hear the message.
- **The database descriptor** — This is used to send summary information to neighbors to synchronize topology databases.
- **The LSR** — This is a request for more detailed information, which is sent when the router receives a database descriptor that contains new information.

- **The LSU** —This is the LSA packet issued in response to the request for database information in the LSR packet.
- **The link-state acknowledgement** —This acknowledges the LSU.

5. *In creating an adjacency, what is the exstart state?*

The exstart state is a stage in the forming of an adjacency between neighbors. This stage is the stage when the DR and the BDR have been elected. The master/slave relationship has been established, as has the initial sequence number of the DDP packets.

6. *What is the database descriptor and when is it used?*

Referred to as DBDs or database descriptor packets (DDPs), these are packets exchanged between neighbors during the exchange state. The DDPs contain summary information taken from the LSAs, which describe the links of every router in the neighbor’s topology table.

7. *Explain the difference between an LSR and an LSA.*

A link-state advertisement (LSA) is a packet describing a router’s links and the state of those links. There are different types of LSAs to describe the different types of links.

An LSR is a link-state request, which is used when the router receives a DDP complete with summary information taken from the LSA. It compares the LSA against the topological database. If either the LSA entry is not present or the entry is older than the DDP, it will request further information via an LSR.

8. *What packet is used to maintain the neighbor table?*

The hello packet is used to maintain the neighbor table. Whenever a hello is heard, the source address in the hello packet is used to reset the hello interval timer. This shows that the neighbor is still active.

9. *What is the metric used by OSPF standards? Is this the same metric that Cisco uses?*

OSPF defines cost as the OSPF metric, but does not define what cost represents. Thus, any determinant could be used and defined manually as cost. Cisco has set a default metric to be the inverse of bandwidth, making the fastest link the most preferred link. This default can be overridden by manual configuration.

10. *Explain the meaning of the letters BDR.*

BDR stands for *backup designated router*. This router acts as the backup to the DR in case the DR fails. The BDR performs none of the DR functions while the DR is operating correctly.

11. *What is used to elect the DR when the election is dynamic?*

When selected dynamically, the DR is elected arbitrarily. The election is made on the basis of the highest router ID or IP address present on the network segment. It is wise to be aware that the highest IP address is the numerically highest number, not the class ranking of the addresses. Therefore, a remote, small router with a Class C address might end up as a DR.

12. *When a new router joins the OSPF network, will it learn about the rest of the OSPF network through the flooding method or the exchange method?*

When a new router connects to a network, it will find a neighbor using the Hello protocol and will exchange routing information.

13. *If an LSA is received that is present in the OSPF database, and the receiving LSA is older than the one currently held by the router, what action is taken?*

The receiving router will send a copy of the LSA it holds in its database to the source of the old LSA and then discard the old LSA it received.

14. *A router has made a neighbor relationship with another router and exchanged DDP. Having compared the routing information from its neighbor, the router realizes that its topology database is incomplete. Name the different stages or states that a router goes through to update its topology database.*

When the router has received the DDPs from the neighboring router, it compares the received network information with that in its topology table. In the case of a new router, such as the 2500, all the DDPs are new. Remember that the DDPs are simply a summary of the routes about which the neighbor knows. If there is a discrepancy between the information in the received DDPs and the router's topology database, then the router requests more detailed information from its neighbor on those routes of which it was unaware. The different stages or states that the router goes through gathering routing information to update the topology database from a neighbor are shown in the following list:

- **The loading state** — If the receiving router, the 2500, requires more information, it will request that particular link in more detail using the LSR packet.

The LSR will prompt the master router to send the LSU packet. This is the same as an LSA used to flood the network with routing information. While the 2500 is awaiting the LSUs from its neighbor, it is in the loading state.

- **The full state** — When these LSRs are received and the databases are updated and synchronized, the neighbors are fully adjacent.

15. *How many equal-cost paths will Cisco enter into the routing table?*

The RFC 2328 that defines OSPF does not state the number of equal-cost paths that can be entered into the routing table. Cisco has defined this to be four paths by default, which can be configured to contain up to six equal-cost paths.

16. *An LSA is received by a router, and when checked against the topology database, it finds the LSA is new or a change in the status of an existing route that has been received. What action will the receiving router take?*

The LSA is flooded out of all the interfaces, excepting the interface through which it was received. The LSA is copied into the topology database, replacing the original LSA if it existed. The received LSA is acknowledged. The SPF algorithm is run to update the routing table.

17. *Which NBMA configuration options are Cisco-specific?*

The configuration options proprietary to Cisco for NBMA are:

- Point-to-multipoint nonbroadcast
- Broadcast
- Point-to-point

18. *What is the difference between a point-to-point interface and a point-to-multipoint interface?*

In a point-to-point network, the concept of broadcast is not relevant because the communication is direct to another router. There is very little network overhead. An IP subnet is required for each point-to-point link.

In point-to-multipoint connections, OSPF simulates a broadcast, the network traffic is replicated and sent down each physical link and uses multicast addressing.

19. *What is the default network type for serial interfaces with HDLC encapsulation, and how often is the hello packet sent?*

The default network type for serial interfaces with HDLC encapsulation is point-to-point and the hello packet is sent out every 10 seconds.

20. *On a multiaccess link, what role does the BDR play?*

The BDR listens to all the OSPF network traffic, which is addressed to both the designated and BDRs. All the routers on the medium have an adjacency with both DRs. The difference is that the BDR listens but does not respond. If the DR fails, the BDR becomes the DR.

Chapter 7

“Do I Know This Already?” Quiz

1. *d*
2. *c, d*
3. *b, d*
4. *b*
5. *c*
6. *b*
7. *d*
8. *b*
9. *c*
10. *b*
11. *b, c*
12. *a, d*
13. *b, c, d*
14. *b*

Q&A

1. *What command is used to determine manually which router on a LAN will become the DR?*

The **priority** command is used to determine manually the DR. The higher the priority, the greater the likelihood is of success. Remember that the default=1 and p=0 means that the router cannot win.

2. *What parameter is used to calculate the default metric of a route in OSPF on a Cisco router?*

The bandwidth parameter configured on an interface is used to determine the default cost or the value of the path with the lowest cost.

3. *It is possible to have more than one OSPF process on a router. How would you do this?*

The **router** command creates the OSPF process with an ID number to identify it. To create another process on the same router, issue the same command again with a different ID number.

It is possible to have more than one process, although it is rarely configured. The process ID in the command **router ospf process-id** not only starts the process, but also identifies the process; repeating the command with another ID number will create another process. One possible scenario for this configuration is a service provider that wants to separate its OSPF domain from its customer.

4. Explain the command `ip ospf network non-broadcast` .

The **ip ospf network non-broadcast** command is the RFC-compliant mode for NBMA. It is the default mode for interfaces and point-to-multipoint subinterfaces. It is used in a full or partial meshed network, and OSPF operates as if on a nonbroadcast network. It is necessary to define manually the DR to be a hub router that is connected to all the other routers. Neighbors must be defined manually.

5. In which of the NBMA configuration choices is it necessary to state the neighbors manually? Why is this necessary?

It is necessary to manually configure the neighbors in the industry-standard NBMA mode and in the Cisco point-to-multipoint nonbroadcast mode.

You need to define the neighbors to the router because the router believes that it is a nonbroadcast medium, so it cannot send out the multicast traffic to ascertain the neighbors.

6. In a Frame Relay environment, which is fully meshed, which OSPF configurations might be chosen? Give reasons for your choice.

The industry-standard NBMA configuration can be chosen in a fully meshed environment. It requires an additional manual configuration of the neighbors, but the network will elect the DR and the BDR. There might be some design concerns about running this mode in an unstable network, which could burden the CPU and the WAN links.

It is possible to use point-to-point subinterfaces without worrying about the OSPF network type because they will become neighbors.

The other alternative is the Cisco broadcast mode, which does not require the manual configuration of neighbors.

7. The Cisco solution point-to-point mode does not require the configuration of DR and BDR. Explain briefly why.

The Cisco solution point-to-point does not require the election of either a DR or a BDR because there are only two nodes on the network. They form an adjacency immediately.

8. *The address 192.100.56.10/21 has been allocated to an interface on the router. This interface alone is to be included in the OSPF process. State the command that would start the process on this interface.*

There are several ways to configure the process to include the interface. The command **network network-number wildcard-mask area area-number** would be a subcommand to the global command **router ospf process-id**. The **network** command is used in both possible solutions; the difference is in the wildcard mask.

- **network 192.100.56.10 0.0.0.0 area 2** — This will match every bit in the interface address.
- **network 192.100.56.100.0.7.255 area 2** — This will also match the interface because it will resolve to the subnet assigned to the wire connected to the interface. This bit allocation was chosen merely to demonstrate the technique. The allocation assumed is the subnet mask of 255.255.248.0. Note that the wildcard mask is the inverse of the subnet mask, ensuring that the individual subnet is selected for the interface.

9. *The metric used by OSPF is cost. How would you change the metric on an interface?*

Underneath the appropriate interface, issue the command **ip ospf cost**. The value for cost is an unsigned integer value expressed as the link-state metric. It can be a value in the range 1 to 65,535.

10. *If the command **ip ospf network non-broadcast** is used, what additional statement is necessary?*

If the command **ip ospf network non-broadcast** is used, the additional statement that is required is the **neighbor** statement. Because the network is a nonbroadcast network that cannot see its neighbors, the neighbors are to be manually configured.

11. *What command shows which router on a LAN is the BDR?*

The **show ip ospf neighbor** command will show the DR and the backup router. Another command that will show the DRs is the **show ip ospf interface** command.

12. *Explain briefly what **show ip ospf database** will reveal.*

The command **show ip ospf database** shows the contents of the topology database and gives a status on the LSAs that have been sent and received, including how long it has been since the last LSA was received.

13. *What command is used to show the state of adjacencies?*

The command **show ip ospf interface** shows the adjacencies that exist with neighbors.

14. *Which command is used to show OSPF packets being sent and received in real time?*

The command **debug ip packet** shows OSPF packets being sent and received in real time.

15. *How would you show the OSPF process ID of the router?*

The commands **show ip ospf**, **show ip ospf database**, and **show ip ospf interface** all show the OSPF process ID on the router.

16. *What makes **debug** a dangerous command for your router?*

The **debug** command has the highest process priority and is therefore capable of consuming all the resources on the router, thus becoming the problem as opposed to helping to solve the problem.

17. *What is the sequence number and where is it held?*

The sequence number is used to ensure the LSA that has been received contains the most recent information about the network. This prevents any packets arriving out of sequence from resulting in a change in the network that is incorrect.

18. *In the **show ip ospf** command, there is a field called the SPF schedule delay. What is the purpose of this field, and what is the default time?*

The SPF schedule delay is the time between OSPF receiving a topology change and starting an SPF calculation. The delay can be an integer from 0 to 65,535. The default time is 5 seconds. If the value is set to 0, this means that the SPF calculation is started as soon as a valid LSA is received.

There is a balance between responding to a topology change quickly and the use of CPU processing.

19. *What is the advantage of the command **show ip ospf interface** in troubleshooting?*

The **show ip ospf** interface command shows how the interface has been configured for OSPF. This allows for the immediate identification of typing errors that result in a mismatch between neighbors.

Chapter 8

“Do I Know This Already?” Quiz

1. *b*
2. *d*
3. *b, d*

4. *c*
5. *b*
6. *d*
7. *d*
8. *a, b, c*
9. *a, c, d*
10. *a*
11. *b*
12. *a, c*

Q&A

1. *In a totally stubby area, which routes are not propagated into the area?*

There will be no summary or external routes propagated by the ABR into the area. Thus, there will be no Type 3, Type 4, or Type 5 LSAs sent into the other area.

2. *Can a virtual link contain a stub area?*

No, a virtual link cannot contain a stub area. A stub area cannot accept external LSAs, and by definition, the virtual link is traversing a different area.

3. *An ABR must be resident in which area?*

An ABR must be resident in Area 0, as well as in the area that is connecting to the backbone area. It has two topological databases, one for each area in which it is resident, so that it knows how to forward traffic.

4. *What LSAs will the ABR forward?*

ABR forwards summary LSAs. It forwards both Type 3 LSAs and Type 4 LSAs. Type 3 LSAs are forwarded to the other ABRs, and Type 4 LSAs are forwarded to the ASBRs. ABR also forwards Type 3 LSAs from other areas into its own area. If the ABR has multiple links in the same area, it also forwards Type 1 and Type 2 LSAs in its capacity as an internal router.

5. *State two advantages in creating areas in OSPF.*

The advantages in creating areas in OSPF include the following:

- It is easier to manage and administrate smaller areas, where hopefully many of the design considerations and even configuration are standardized.

- It uses a smaller topology table, which reduces the CPU, memory, and network bandwidth consumption.
- Fewer SPF calculations are involved because the topology table is smaller and there is less likelihood of change in the network.
- It uses a smaller routing table if summarization is in operation.

6. *What is an external route, and on which type of router will this route be introduced?*

An external route is a route that did not originate in the OSPF domain. It has been redistributed from another routing protocol or static routing. An external route is introduced into the OSPF domain by an ASBR.

7. *Why is the use of summarization important in the design of OSPF?*

Summarization is important in the design of OSPF because it supports a hierarchical design and allows for the summarization of IP subnets between areas, which reduces the size of the routing tables, which in turn reduces the CPU and memory requirements.

8. *How many routers does Cisco suggest is the limit to have in a single area?*

Cisco suggests that there should not be more than 50 routers in any single area.

9. *What are the restrictions to be considered in the creation of a stub area or a totally stubby area?*

Some restrictions govern creating a stub or totally stubby area. Because no external routes are allowed in these areas, the following restrictions are in place:

- No external routes are allowed.
- No virtual links are allowed.
- No redistribution is allowed.
- No ASBR routers are allowed.
- The area is not the backbone area.
- All the routers are configured to be stub routers.

10. *A virtual link in OSPF is used to solve what problem?*

The virtual link provides a disconnected area with a logical path to the backbone. The virtual link must be established between two ABRs that have a common area, with one ABR connected to the backbone. It can also be used to connect two area 0s together. This might be necessary

when two companies merge, each with its own area 0, or if, due to the loss of a link, the area 0 becomes bisected.

11. *State one disadvantage for making an NBMA cloud Area 0.*

Creating the NBMA cloud as one OSPF area, preferably Area 0, causes summary LSAs to be flooded throughout the NBMA network. This results in a large number of routers recalculating whenever there is a change that requires the topology table to be updated, and the Frame Relay cloud can become saturated. If the Frame Relay cloud has a problem, then the entire network might suffer.

12. *State one advantage in making the centralized routers and network resources dwell in Area 0 while the Frame Relay cloud and the stub remote LANs reside in satellite stub areas.*

One advantage of this design is that any flooding of external LSAs is prevented from entering the Frame Relay network, because it is a stub network. This reduces the network overhead.

13. *How does creating a number of areas in OSPF reduce the number of SPF calculations?*

The number of SPF calculations is reduced because the size of the topology table is reduced. This lessens the likelihood of a change in the network and, thus, SPF calculations.

14. *How does a stub area differ from the backbone area?*

A stub area differs from the backbone area in that it does not propagate external routes into its area. The backbone is obliged to forward these LSAs to ensure connectivity throughout the network.

15. *How does a totally stubby area differ from a stub area?*

A totally stubby area differs from a stub area in that it propagates neither external routes nor summary routes from other areas. This is a Cisco solution to minimize the amount of CPU and memory required of the routers within the area. Connectivity is achieved by the use of default routes, which are advertised to the internal routers.

16. *State the different LSA types.*

The different LSA update types are as follows:

- **Router link** — Sent by the router, stating the links directly connected. These are flooded through the area. This update is identified by the type code Type 1.
- **Network link** — Sent by the DR, stating the links for the LAN for which it is the DR. These LSAs are flooded throughout the area. This update is identified by the type code Type 2.

- **Summary link** — Sent by the ABR into the backbone. It states the IP subnets within the area that are to be advertised into other areas. This is where summarization would be configured. This update is identified by the type code Type 3.
- **Summary link (to an ASBR)** — Sent from an ABR to a router that connects to the outside world (ASBR). It contains the metric cost from the ABR to the ASBR. This update is identified by the type code Type 4.
- **External link** — Sent to the ASBRs to which the organization is directly connected. This update is identified by the type code Type 5.
- **The NSSA External LSA** — These LSAs are created by the ASBR residing in an NSSA. This LSA is similar to an autonomous system external LSA, except that this LSA is contained within the NSSA area and is not propagated into other areas.

17. *Where does the backbone router reside, and what is its function?*

OSPF has special restrictions when multiple areas are involved. If more than one area is configured, one of these areas must be Area 0. This is called the backbone. When designing networks, it is good practice to start with Area 0 and then expand into other areas later.

The backbone must be at the center of all other areas—that is, all areas must be physically connected to the backbone. The reasoning behind this is that OSPF expects all areas to inject routing information into the backbone; in turn, the backbone will disseminate that information into other areas.

18. *There are two types of summarization. What are they?*

The two types of summarization are as follows:

- **Interarea route summarization** — These routes are sent between areas. The ABR will summarize routes if the network within the area was designed using contiguous addresses, conforming to both a physical and a logical hierarchy.
- **External route summarization** — These are routes sent into OSPF from another routing protocol. This summarization also demands a hierarchical design using contiguous addresses. This is employed at the ASBR.

19. *For how many LANS does Cisco suggest a router should serve as a DR or a BDR?*

Cisco suggests that a router should be a DR or a BDR for only one LAN.

20. *Which router type creates LSA Types 3 and 4?*

The ABR creates the LSA Types 3 and 4.

Chapter 9

“Do I Know This Already?” Quiz

1. *a*
2. *b*
3. *a, d*
4. *a, c*
5. *b*
6. *a*
7. *a*
8. *d*
9. *b*
10. *a, c, d*
11. *b, c*
12. *a, b, d*
13. *c*

Q&A

1. *Which command in OSPF shows the network LSA information?*

The command **show ip ospf** [*process-id area-id*] **database network** displays the network link-state information.

2. *What command would you use to create a totally stubby area?*

The command **area area-id stub no-summary** will create a totally stubby area. This is a subcommand to the **router ospf process-id** command. It is necessary only on the ABR, but all the other routers in the area must be configured as stub routers.

3. *What is a virtual link, and what command would you use to create it?*

A virtual link is a link that creates a tunnel through an area to the backbone (Area 0). This allows an area that cannot connect directly to the backbone to do so virtually. The command to create the link is **area area-id virtual-link router-id**. Note that the *area-id* that is supplied is that of the transit area, and the *router-id* is that of the router at the other end of the link. The command needs to be configured at both ends of the tunnel.

4. *Where would you issue the command to summarize IP subnets? State the command that is used.*

Summarization is done at area boundaries. The command to start summarization is the **area range** command, with the syntax **area area-id range address mask**. To summarize external routes, use the **summary-address** command on the ASBRs.

5. *How would you summarize external routes before injecting them into the OSPF domain?*

The command **summary-address address mask** is the command that you would use.

6. *When is a virtual link used?*

A virtual link is used when an area is not directly attached to the backbone area (Area 0). This may be due to poor design and a lack of understanding about the operation of OSPF, or it may be due to a link failure. The most common cause of an area separating from the backbone is link failure, which can also cause the backbone to be segmented. The virtual link is used in these instances to join the two backbone areas together. Segmented backbone areas might also be the result of two companies merging.

7. *Give the command for defining the cost of a default route propagated into an area.*

The command to define the cost of a default route propagated into another area is **area area-id default-cost cost**.

8. *Give an example of when it would be appropriate to define a default cost.*

It is appropriate to define a default cost for the default route when a stub area has more than one ABR. This command allows the ABR or exit point for the area to be determined by the network administrator. If this link or the ABR fails, the other ABR will become the exit point for the area.

9. *On which router is the area default cost defined?*

The default cost for the default route is defined on the ABR. The ABR will then automatically generate and advertise the route cost along with the default route.

10. *Give the command to configure a stub area and state on which router it is configured.*

The command syntax to configure a stub area is **area area-id stub**. This command is configured on the ABR connecting to the area and on all the routers within the area. Once the configuration is completed, the Hellos are generated with the E bit set to 0. All routers in the area will only form adjacencies with other routers that have the E bit set.

11. *What is the purpose of the **area range** command, and why is it configured on the ABR?*

The **area range** command is configured on an ABR because it dictates the networks that will be advertised out of the area. It is used to consolidate and summarize the routes at an area boundary.

12. *Give the commands to configure a router to place subnets 144.111.248.0 through to 144.111.255.0 in Area 1 and to put all other interfaces into Area 0.*

The commands are as follows:

```
network 144.111.248.0 0.0.7.255 area 1
network 0.0.0.0 255.255.255.255 area 0
```

13. *Give the syntax to summarize the subnets 144.111.248.0 to 144.111.254.255 into another autonomous system.*

The syntax is as follows:

```
summary-address 144.111.248.0 255.255.248.0
```

14. *Explain briefly the difference between the **area range** command and the **summary-address** command.*

The **area range** command is used to summarize networks between areas and is configured on the ABR. The **summary-address** command is used to summarize networks between autonomous systems and is configured on the ASBR.

15. *Explain the following syntax and what it will achieve: **area 1 stub no-summary** .*

The command **area 1 stub no-summary** creates a totally stubby area. The number after the word *area* indicates the area that is being defined as a totally stubby area. This is necessary because the router might be an ABR with connections to many areas. Once this command is issued, it prevents summarized and external routes from being propagated by the ABR into the area. To reach the networks and hosts outside the area, routers must use the default route advertised by the ABR into the area.

16. *Why would you configure the routing process to log adjacency changes as opposed to turning on debug for the same trigger?*

The reason to configure the router process to log adjacency changes to syslog as opposed to running debug is an issue of resources. It takes fewer router and administrator resources to report on a change of state as it happens than to have the debugger running constantly. The debug process has the highest priority and thus everything waits for it.

17. *Give some of the common reasons that neighbors fail to form an adjacency.*

Many OSPF problems stem from adjacency problems that propagate throughout the network. Many problems are often traced back to neighbor discrepancies.

If a router configured for OSPF routing is not seeing an OSPF neighbor on an attached network, do the following:

- Make sure that both routers are configured with the same IP mask, MTU, Interface Hello timer, OSPF Hello interval, and OSPF dead interval.
- Make sure that both neighbors are part of the same area and area type.
- Use the **debug** and **show** commands to trace the problem.

18. *When configuring a virtual link, which routers are configured?*

The configuration is between the ABRs, where one of the ABRs resides in Area 0 and the other in the area that is disconnected from the backbone. Both of the ABRs are also members of the transit area. Having created the virtual link, both ABRs are now members of Area 0, the disconnected area, and the transit area.

19. *What does the command **area 1 default-cost 15** achieve?*

The command **area 1 default-cost 15** will assign a cost of 15 to the default route that is to be propagated into the stub area. This command is configured on the ABR attached to the stub area.

20. *Explain what is placed in the parameters **area-id** and **router-id** for the command **area area-id virtual-link router-id**.*

The parameter *area-id* is the area ID of the transit area. So if the ABR in Area 0 is creating a virtual link with the ABR in Area 3 through Area 2, the area ID stated in the command is Area 2. The router ID is the router ID of the router with whom the link is to be formed and a neighbor relationship and adjacency established.

Chapter 10

“Do I Know This Already?” Quiz

1. *b*
2. *d*

3. *a*
4. *a, b, c, d*
5. *b*
6. *d*
7. *b*
8. *a*
9. *c*
10. *d*
11. *b*
12. *d*
13. *a*
14. *a, b*
15. *b, c*
16. *a, b*
17. *a*
18. *b*
19. *a, b, d*
20. *a, b, d*
21. *c*

Q&A

1. *Which system generates the pseudonode?*

The DIS generates the pseudonode, using its own system ID and setting the following octet to be a nonzero number.

2. *What is a CSNP? When is it used?*

The complete sequence number packet describes every link in the link-state database. It is sent on point-to-point links when the link comes up to synchronize the link-state databases. The DIS on a multicast network will send out CSNPs every 10 seconds.

3. *What is a PSNP? When is it used?*

Partial Sequence Number Packets are sent on point-to-point links to explicitly acknowledge each LSP they receive. A router on a broadcast subnetwork will send a PSNP requesting the LSPs it needs to synchronize its link-state database.

4. *A new router comes online on a multiaccess link, and the priority is the same as that of the DIS for the segment. What action is taken?*

The routers exchange Hellos and immediately see that both routers have the same priority. If the new router had a higher priority, it would take over as the new DIS. However, if both routers have the same priority, the router with the highest MAC address will reign as the DIS.

5. *What happens in the event of the DIS dying?*

There is no backup designated router in IS-IS. Therefore, if the DIS meets an untimely death, a new DIS would be elected, based on priority or highest MAC address. If another router comes online with a higher priority, it will dislodge the existing DIS and rule in its place. This behavior is different from that of OSPF. Once a new DIS is elected, the link-state databases are purged and new LSPs are flooded.

6. *Integrated IS-IS can be used to send information about which routed protocols?*

IS-IS is capable of carrying both IP and CLNS.

7. *How often does the DIS send out a Hello packet?*

The DIS sends out hellos every 3.3 seconds, three times the speed of other routers on the multiaccess link.

8. *What is the name of the link-state algorithm used to create entries for the routing table?*

The name of the link-state algorithm is the Dijkstra algorithm.

9. *What is the relationship between the Hello timer and when the path is considered to have died?*

The default timer is three times that of the Hello timer; thus the path will wait for 30 seconds before declaring the path dead and flushing the LSPs from the link state database.

10. *Integrated IS-IS areas are similar to which type of areas in OSPF?*

Integrated IS-IS areas are similar to OSPF stub areas.

11. *Describe one design restriction in configuring Level 2 routing.*

There is only one hard and fast rule for the design of a Level 2 network: Level 2 routers must be contiguous; that is, the area cannot be fractured.

12. *Given the following address:*

49.0001.2222.2222.2222.00

Is this a NET or NSAP address? Give reasons for your choice.

The address is a NET address because the last octet is set to 0x00. This there is no network service defined. This is the address of a router, not an end system.

13. *What is a pseudonode and what is its purpose?*

The pseudonode is the LAN identifier for a broadcast subnetwork. The pseudonode is the System ID of the DIS plus the Circuit ID. The pseudonode has links to each of the ISs, and each IS has a single link to the pseudonode. The use of the pseudonode reduces the number of links required. Instead of $n-1$ links to each of the other ISs, there is one link per IS. The DIS generates link-state PDUs on behalf of the pseudonode. These LSPs are sent to all the connected ISs.

14. *State two reasons why a router might not be able to find a neighbor.*

For an adjacency to be formed and maintained, both interfaces must agree on the following:

- The same MTU.
- Both are Level 1. If both are Level 1, they must be in the same area.
- Both are Level 2.
- At least one is Level 1-2.
- The authentication must be the same.
- The Hello timers (including the holddown timer) must match. If one router has a Hello timer of 40 seconds, the defaults on the other router would time out the holddown timer and purge the LSP, resulting in a flapping link and endless SPF calculations.

15. *Explain briefly why two routers cannot have the same system ID within the area.*

The system ID is the unique identifier for the area. The first part of the address is a very long area address, of which only the last six octets are available for addressing the router or host.

16. *What does TLV stand for? Briefly explain its purpose.*

TLV is the same as CLV, but some literature refers to the variable length fields as Type/Length/Value in accordance to the IP terminology. Although the IS-IS PDUs are fixed, the TLV fields are variable length and can expand as needed. This design allows great flexibility and movement to develop in step with technological advances. The development of TLV code 128 extended IS to carry integrated IS-IS.

17. *How many link-state databases does a Level 1-2 router have?*

A Level 1-2 router has two link-state databases, one for the Level 1 routes and the other for the Level 2 routes. A separate SPF algorithm is run for each database.

18. *Integrated IS-IS packets run directly on top of which layer?*

IS-IS packets run directly on top of the data-link layer.

19. *What is the NET address associated with in the IS-IS addressing scheme?*

The NET address is associated with the end system, but not with a process on the end system. The address is that of an entire system, as opposed to an interface on the system, as is the case with IP. Because the NET (unlike the NSAP) does not identify a process, the address is that of a transitional or intermediate system. Therefore, the NET address is associated with the router or IS and is the destination address of the next hop in the life of a routed or routing packet.

20. *Describe briefly the main characteristics of a Level 1 IS.*

Characteristics of a Level 1 IS include the following:

- An intra-area router.
- Similarity to an OSPF stub router.
- Knowledge of the network limited to the area.
- A link-state database with all the routing information for the area.
- The address of the closest Level 2 router to which to send traffic destined for another area.
- Neighbors must be in the same area.
- A DIS is elected on LANs.

Chapter 11

“Do I Know This Already?” Quiz

1. *a, b, c*
2. *b*
3. *b, c*
4. *a*
5. *c*
6. *b*
7. *c*
8. *b*
9. *d*
10. *b, c*
11. *a*
12. *c*

Q&A

1. *Which systems would you configure as Level 1-2 systems?*

It is necessary to configure routers that straddle more than one area as Level 1-2 routers so that they can receive updates from both Level 1 and Level 2 routers and thus forward datagrams from Level 1 routers out of their area. Some designs allow for every router to be configured as a Level 1-2 router; this is the default configuration on Cisco routers. This eliminates errors but is heavy on network resources.

2. *Which IS-IS configuration uses a full mesh and simulates a broadcast technology?*

The multipoint configuration option is used with a full mesh and is seen by IS-IS as a broadcast network. It will therefore elect a DIS for the network.

3. *What are the four stages of the routing process?*

The four stages of the routing process are update, decision, forwarding, and receive.

4. *What does an LSP contain?*

An LSP contains the list of neighbors connected to the originating router.

5. *When are LSPs generated?*

LSPs are generated whenever there is a change in the network, often because of a configuration change. However, any of the following instances trigger a new LSP to be flooded throughout the network:

- An adjacency comes either up or down (for example, a new router comes online).
- An interface on the router changes state or is assigned a new metric.
- An IP route changes (for example, because of redistribution).

6. *State at least one of the main steps of the flooding process on a point-to-point link.*

The following list describes the flooding process on a point-to-point link:

- a. When an adjacency is established, both sides send a CSNP packet with a compressed version of their link-state databases.
- b. If there are any LSPs in the receiving router’s database that were not in the CSNP it received, it sends a copy of the missing LSPs to the other router.
- c. Likewise, if the database is missing any LSPs received in the CSNP, the receiving router requests the detailed LSP to be sent.
- d. The individual LSPs are requested, sent, and acknowledged via PSNPs.
- e. When an LSP is sent, the router sets a timer. If no explicit acknowledgement has been received before the timer expires, the LSP is resent. This timer is the **minimumLSP-Transmission-interval** and can be configured; the default on a Cisco router is 5 seconds.

7. *Which three fields determine whether the LSP is valid?*

The LSP contains three fields that help determine whether the LSP that has been received is more recent than that held in the database and whether it is intact or has been corrupted. These three fields are as follows:

- **Remaining Lifetime** : This is used to age-out old LSPs. If an LSP has been in the database for 20 minutes, it is assumed that the originating router has died. The refresh timer is set to 15 minutes.

If the lifetime expires, the LSP has the content removed, leaving only the header. The lifetime is set to show that it is a new LSP, and then it is flooded through the network. All receiving routers accept the mutilated LSP, recognize that this means the route is bad, and purge the existing LSP from their databases.

- **Sequence Number** : This is an unsigned 32-bit linear number. The first LSP is allocated the sequence number 1, and the following LSPs are incremented by 1.

- **Checksum** : If a router receives an LSP and the checksum does not compute correctly, the LSP is flushed and the lifetime is set to 0. The router floods the LSP, all routers purge the LSP, and the originating router retransmits a new LSP.

8. *Once the link-state databases are synchronized, the Dijkstra algorithm is run. Describe where the router places itself in the tree.*

Each router builds a shortest path tree (SPT) with itself as the root. This is achieved by taking all the LSPs from the link-state database and using the Dijkstra algorithm to create the SPT. The SPT is used in turn to create the forwarding table, which is also known as the routing table.

9. *State two criteria in determining which paths are to be placed in the forwarding database.*

If there is more than one path to a remote destination, the criteria by which the lowest cost paths are selected and placed in the forwarding database are as follows:

- If there is more than one path with the lowest value metric, Cisco equipment places up to six equal-cost paths into the table. The default number of equal-cost paths is four.
- Optional metrics are chosen before the default metric, but because Cisco supports only the default metric, this is a moot point.
- Internal paths are chosen before external paths, because going outside the autonomous system is likely to be a suboptimal route and might be the result of a routing loop.
- Level 1 paths within the area are more attractive. If the path is within the area, not only is it more efficient to route directly to it, but also going outside the area and returning can be the cause of a routing loop, demanding greater resources and time.
- The address with the longest match or most specific address in IP is the address with the longest IP subnet mask. This ensures that the closest router is chosen, because prefix routing is configured by summarization that can occur only on area boundaries.

10. *What are the ISO metrics?*

The metrics defined in ISO 10589 are as follows:

- **Default**: Sometimes referred to as *cost*. Every Integrated IS-IS router must support this metric. Cisco set the default for all interfaces to be 10.
- **Delay**: This optional metric reflects the transit delay.
- **Expense** : This optional metric reflects the monetary expense of the network.
- **Error** : The reliability of the path is determined as the metric.

- 11.** *How many equal-cost paths is it possible to have in the IS-IS routing table of a Cisco router?*

The default number of equal-cost paths allowed in the routing table is four, although Cisco allows six to be placed in the table.

- 12.** *What is a narrow metric?*

A narrow metric is the default metric, which has a 6-bit field. Cisco increased the size of the field to 24 bits.

- 13.** *Where is the IS-IS metric applied?*

The IS-IS metric is applied to the outgoing interface.

- 14.** *What action will the routing process take if it sees an incomplete LSP fragment?*

If an LSP fragment is incomplete, the routing process ignores it, safe in the knowledge that it will be retransmitted if the sending router does not receive an ACK within a specified time frame.

- 15.** *Why is the IS-IS default of cost the only metric supported by Cisco?*

Each metric that is configured for use in IS-IS requires its own database. If the router is a Level 1-2 router, it will need a database for each metric and each level of routing. This could result in eight databases and the use of many resources from both the router and the network.

- 16.** *When designing a network for fast convergence, what is the compromise that you need to consider?*

Typically, the trade-off is between reliability and speed. To increase the speed of convergence for a routing protocol, it might be sufficient to tune the update process, although this results in the compromise of resources and reliability. If you reduce the update timers, the databases converge more quickly, but the network could be depleted of necessary resources to route data.

- 17.** *What is a suboptimal routing decision?*

Suboptimal routing decisions occur when Level 1 areas have knowledge of only networks within their own areas. To reach another area, packets are sent to the nearest Level 2 router. Without additional configuration, the Level 1 router determines the nearest Level 2 router to be the one with the lowest hop count. The metrics used are the default metric of 10 on each outbound interface; therefore, the best route translates to that with the lowest hop count. As you know, the router two hops away might include a 16 Mbps Token Ring and a 56 kbps link as opposed to the three hops of Fast Ethernet and ATM.

18. *Where does route summarization take place?*

Route summarization can be configured on a Level 1-2 router at the area boundary.

19. *When is a DIS elected in a WAN environment?*

A DIS is elected on a WAN when the NBMA cloud is configured as multipoint.

20. *Explain briefly how the IS-IS NBMA cloud is different than the configuration of the OSPF cloud.*

Frame Relay and ATM are examples of NBMA networks, which are not accommodated in Integrated IS-IS. OSPF has a point-to-multipoint configuration option, but Integrated IS-IS does not. The solution in Integrated IS-IS is to configure the link as multipoint, allowing the election of a DIS. The alternative is to configure the interfaces with subinterfaces that are point-to-point.

Chapter 12

“Do I Know This Already?” Quiz

1. *b*
2. *c*
3. *d*
4. *c*
5. *b*
6. *a*
7. *d*
8. *c*
9. *b, c, d*
10. *b*
11. *b, c, d*
12. *c, d*

Q&A

1. *What command is used to configure the Integrated IS-IS router process?*

The **router isis** global configuration command starts the routing process for integrated IS-IS.

2. *What is the default routing level on a Cisco router?*

By default, the routing process for Integrated IS-IS runs as a Level 1-2 router.

3. *What command is used to configure Integrated IS-IS routing on the interface?*

The **ip router isis** interface command is used to start Integrated IS-IS routing on the interface.

4. *How is the NET address configured on the router?*

Under the command that started the routing process for Integrated IS-IS, enter the **net** command followed by the full network address.

5. *What command is used to show the state of adjacencies on the router?*

The **show clns neighbor** and **show clns interface** commands both display the adjacencies. The **show clns neighbor** command gives information as to the state of the link, the type of routing performed on the link, the data-link address (SNPA) of the link, and the holdtime, or how long since it received the last Hello.

6. *What command identifies the designated intermediate system router for your LAN?*

The designated intermediate system (DIS) is identified in the output screen of the **show clns interface** command or any **show** command that shows the pseudonode, such as **show isis database** or **show clns neighbors**. If the medium is multiaccess broadcast, the Circuit ID field shows the ID of the pseudonode. The pseudonode is identified by the nonzero value in the octet following the system ID of the DIS (for example, R2.01).

7. *Explain briefly what **show isis database** reveals.*

The **show isis database** command displays the LSPs in the link-state database. This database should be identical on every router in the area.

8. *What command reveals the trigger for the last SPF calculation on the router?*

The **show isis spf-log** command shows the trigger for the last 20 occurrences for which the SPF calculation was run.

9. *For Frame Relay, when would you configure the **map** command with the broadcast parameter?*

The Frame Relay **map** command with the broadcast parameter is used in a fully meshed environment and when the network is multiaccess. This allows the election of a DIS, streamlining of adjacencies, and the efficient use of IP subnets.

10. *Which command is used to display all update packets that are both received and sent by a router?*

The command that shows all the Integrated IS-IS packets both sent and received by the router is **debug isis update-packets** . These packets are the sequence number PDUs (CSNPs and PSNPs) and LSPs that are detected by the router.

11. *State two reasons why a router may not be able to find a neighbor.*

For an adjacency to be formed and maintained, both interfaces must agree on the following:

- The same MTU must be configured on both interfaces.
- If one of the routers is configured as Level 1, both routers must have Level 1 capability.
- The system IDs must be unique to the router.
- Level 1 routers must be in the same area.
- If the routers are in different areas, at least one of the routers must be Level 1-2.
- The authentication configuration must be the same on both routers.
- The Hello timers (including the holddown timer) must match. If one router has a Hello timer of 40 seconds, the defaults on the other router would time out the holddown timer and purge the LSP, resulting in a flapping link and endless SPF calculations.

12. *Which command shows the LSPs in detail?*

The command **show isis database detail** shows the LSPs in detail.

13. *How would you ensure that an adjacency has been established?*

The commands that display the interface and the adjacency on the local router are either the **show clns neighbor** or the **show clns interface** .

14. *What are the steps required for a basic configuration in IS-IS?*

The steps required for a basic configuration in IS-IS are:

- Define the areas and addresses.
- Enable IS-IS on the router.
- Configure the NET address.
- Enable IS-IS on the appropriate interfaces.

15. *Give the commands required to summarize the networks 10.10.0.0 through to 10.10.255.0 into another area of IS-IS.*

The commands required to summarize the networks 10.10.0.0 through to 10.10.255.0 into another area of IS-IS are:

- **router isis**
- **summary-address 10.10.0.0 255.255.0.0**

16. *Which command verifies the circuit type and the metric?*

The command **show clns interface** shows both the circuit type, or routing level, of the interface and the IS-IS metric for outgoing packets.

17. *Which network topology defaults in the Broadcast mode?*

Essentially, all networks, except true point-to-point topologies, are treated as broadcast networks. That means Ethernet, Token Ring, FDDI, ATM, Frame Relay, and X.25 are all seen as broadcast networks.

18. *For which WAN topology is a point-to-point configuration recommended?*

In an NBMA multipoint configuration, subinterfaces and a point-to-point IS-IS network should be configured. NBMA multipoint topologies should be avoided, because, although they can be made to work, they require complex configuration and do not work with the strengths of the IS-IS routing protocol.

19. *When is it necessary to map CLNS to the DLCI?*

It is necessary to map CLNS to the DLCI in a NBMA Frame Relay network that is using a point-to-multipoint topology.

The **frame-relay map ip** command maps the IP destination address to the outgoing DLCI and defines the interface as a broadcast interface. Integrated IS-IS uses the links as if they were truly a broadcast link and elects a DIS.

The **frame-relay map clns** command maps to the CLNS process on the destination router. Without the second command, no routes appear in the IP routing table because CLNS does not receive the frames to populate the IP routing table. Remember that these are IP routes carried in the IS-IS routing protocol. It is IS-IS that updates the IP routing table.

20. In the **show clns interface** command, it is possible to identify the DIS on a multiaccess link. Which field in the output screen of this command would show the DIS for the segment?

The field that shows the DIS is the circuit ID. This field shows the pseudonode ID, which is the system ID of the DIS. It is easily identifiable, as it has a value greater than 0x00 in the octet after the system ID.

Chapter 13

“Do I Know This Already?” Quiz

1. *c*
2. *a, b, c*
3. *a, d* [The four areas are protocol independence, Reliable Transport Protocol (RTP), neighbor discovery and recovery, and DUAL.]
4. *b, c*
5. *a, d*
6. *b*
7. *c, d*
8. *a*
9. *d*
10. *b*
11. *c*
12. *b*

Q&A

1. If a router does not have a feasible successor, what action will it take?

If the router does not have a feasible successor in its topology table, it sends a query packet to its neighbors asking whether they have a feasible successor.

2. When does EIGRP need to be manually redistributed into another EIGRP process?

EIGRP needs to be manually redistributed into another EIGRP process when the autonomous system number is different.

3. *Which timers are tracked in the neighbor table?*

The timers that the neighbor table keeps track of are the holdtime, the SRTT, and the RTO.

4. *What is the difference between an update and a query?*

An update is the routing information packet that a router will send out to inform its neighbors of a change in the network. In a query, the router has no feasible successor in its topology table for a network that is down. At this point, it queries its neighbors to ascertain whether they have a feasible successor. If they do, this route becomes the feasible successor for the original router.

5. *When does EIGRP recalculate the topology table?*

EIGRP recalculates the topology table whenever it receives a change input to the topology table. This could be a change of metric for a physically connected link; a change of status of a physically connected link; or an EIGRP routing packet, either an update, a query, or a reply packet.

6. *EIGRP has a default limit set on the amount of bandwidth that it can use for EIGRP packets. What is the default percentage limit?*

The default percentage limit of bandwidth allocated to EIGRP packets is 50 percent.

7. *State two rules for designing a scalable EIGRP network.*

The rules for scaling an EIGRP network include the following:

- Allocation of addresses should be contiguous to allow summarization.
- A hierarchical tiered network design should be used to allow summarization.
- Sufficient network resources (both hardware and software) should be used on network devices.
- Sufficient bandwidth should be used on WAN links.
- Appropriate EIGRP configuration should be used on WAN links.
- Filters should be used.
- Network monitoring should be used.

8. *EIGRP can be used to send information about which three routed protocols?*

EIGRP can be used as a routing protocol for IP, IPX, and AppleTalk.

9. *Which EIGRP packets are sent reliably?*

The packets that EIGRP sends reliably are updates, queries, and replies. EIGRP uses RTP. This is necessary because EIGRP does not send out periodic updates, and the RTP mechanism ensures a loop-free synchronized network.

10. *In what instances will EIGRP automatically redistribute?*

EIGRP will automatically redistribute between itself and IGRP as long as both processes are running the same autonomous system number.

EIGRP for IPX automatically redistributes into IPX for RIP/SAP and EIGRP for AppleTalk; it similarly redistributes automatically into RTMP.

11. *How long is the holdtime, by default?*

The holdtime is three times the Hello timer. The holdtime is 15 seconds or 180 seconds, depending on the medium.

12. *What is an EIGRP topology table, and what does it contain?*

The topology table contains all links received with a metric other than infinity, that is, every valid path. This will not be all links, because of split horizon. The metric for every path is held along with the metric from the next logical hop or neighbor. The table contains the outgoing interface on the router through which to reach the remote network and the IP address of the next-hop address. The status of the route (passive or active) is also recorded. The topology table also keeps track of the routing protocol packets that have been sent to the neighbors.

13. *What is the advertised distance in EIGRP, and how is it distinguished from the feasible distance?*

Advertised distance is the metric that is reported by the neighbor routers. Feasible distance is the metric that is reported by neighbor routers, plus the cost associated with the forwarding link from the local interface to the neighbor routers.

14. *What EIGRP algorithm is run to create entries for the routing table?*

The Diffusing Update Algorithm (DUAL) is run on the topology table. It is used to determine the best path and to build the routing table.

15. *When does EIGRP place a network in active mode?*

EIGRP places a network into active mode when there is no feasible successor in its topology table.

16. *By default, EIGRP summarizes at which boundary?*

By default, EIGRP summarizes at autonomous system boundary and at the classful network boundary.

17. *What is Stuck in Active?*

Stuck in Active (SIA) is a state in which a router will place a route after it has failed to hear a reply to a query that was sent to a neighbor.

EIGRP sends a query when a route is lost and another feasible route does not exist in the topology table. The SIA is caused by two sequential events: First, a route has gone away. Second, an EIGRP neighbor (or neighbors) has not replied to the query for that route. When the SIA occurs, the router clears the neighbor that has not replied to the query. When this happens, it is necessary to determine which neighbor has been cleared, keeping in mind that this router could be many hops away.

18. *State two factors that influence EIGRP scalability.*

A hierarchical tiered design and contiguous addressing are both critical to being able to scale an EIGRP network. If these are in place, it is possible to summarize the network, which reduces the network resources needed for large tables and limits the query range of the router. It is also important to ensure that the router has sufficient memory, the network has sufficient bandwidth on its WAN links, and, where appropriate, the bandwidth command has been configured.

19. *What are reply packets in EIGRP?*

The reply packet is used to update the topology table. It is a response to a query sent out by a neighbor asking about suspect routes.

20. *What conditions must be met for a router to become a neighbor?*

To become a neighbor, the following conditions must be met:

- The router must hear a Hello packet or an ACK from a neighbor.
- The autonomous system number in the packet header must be the same as that of the receiving router.
- The neighbor’s k-values and metric settings must be the same as that of the receiving router.

Chapter 14

“Do I Know This Already?” Quiz

1. *b*
2. *a*
3. *c*
4. *d*
5. *b*
6. *c*
7. *d*
8. *b*
9. *a, b, c, d*
10. *c*
11. *a*
12. *c*

Q&A

1. Which command has superseded the **passive-interface** command for preventing EIGRP traffic from traversing a link?

The **network** *network wildcard-mask* suppresses Hellos between interfaces not included in the **network** command. This means that no routing updates can be sent or received on that interface.

2. What is the preferred configuration for a hybrid multipoint NBMA network when one VC has a CIR of 56 kbps and the other five VCs each have a CIR of 256 kbps?

The preferred configuration for a hybrid multipoint NBMA network, in which one circuit is lower than the other circuits, is to create a point-to-point subinterface for the lower circuit and then to configure the bandwidth to reflect the CIR of the link. Another subinterface should be created as a multipoint interface with a configured bandwidth that equals the aggregate CIR of all the circuits—in this instance, $5 * 256$ kbps, or 1280 kbps. An alternative solution is to configure each PVC as a point-to-point link.

3. *With four Frame Relay circuits in a multipoint solution and a bandwidth configuration of 224, what is the EIGRP bandwidth allocation per circuit, and where would the **bandwidth** command be configured?*

The command would be configured on the physical interface. The CIR of each circuit is 56 kbps.

4. *Explain the purpose of the command **no auto-summary**.*

The command **no auto-summary** is used to turn off automatic summarization, which in EIGRP happens at the IANA or major network boundary. This command is used in conjunction with the interface commands for manual summarization.

5. *Explain the meaning of the command **ip bandwidth-percent eigrp 63 100**.*

This command overrides the default bandwidth of 50 percent that is allocated to EIGRP for network overhead. This command sets the bandwidth allocation to be 100 percent of the link for the EIGRP autonomous system of 63 on the interface upon which it is configured. This command would be used if the **bandwidth** command had set the logical bandwidth of the link to be artificially low.

6. *In what instances will EIGRP automatically redistribute?*

EIGRP will automatically redistribute between other EIGRP processes with the same autonomous system. EIGRP will also redistribute between itself and IGRP as long as both processes are running the same autonomous system number.

7. *How long is the holdtime, by default?*

The holdtime is three times the Hello timer. The holdtime is 15 seconds or 180 seconds, depending on the medium.

8. *For what is the **variance** command used?*

The **variance** command is used to determine additional paths to be included in load balancing traffic to remote networks. The command is used in conjunction with the multiplier number. This number multiplies the path with the best (lowest) metric by the number stated as the multiplier. Any paths that the router knows of that have a metric value equal to or less than this value are included in the paths for load balancing. The amount of traffic sent across each path will be proportional to the metric value of the path.

9. *What command is used to display the passive and active state of the routes?*

The command **show ip eigrp topology** shows the passive and active state of the routes that are contained in the table. The route is passive if the route is operational; it is set in an active state

if the router must query its neighbors for alternative paths to a network. This command also shows the number of successors and the neighbors and distance information.

10. *What command is used in EIGRP to perform manual summarization?*

The interface command **ip summary-address autonomous-system-number address mask** is used to define the summary address to be used.

11. *For Frame Relay, when would you configure the physical interface (as opposed to a subinterface) with the **bandwidth** command?*

If all the circuits have the same CIR, the **bandwidth** command can be used on the physical interface. The interface will divide the available bandwidth set on the command by the number of circuits.

12. *Which command is used to display all types of EIGRP packets that are both received and sent by a router?*

The command **debug eigrp packet** displays the types of EIGRP packets that are both sent and received by the router.

13. *What problems can be solved with the configuration of summarization?*

Summarization reduces the amount of resources needed by both the network and the routers within the network. The routing tables are reduced, which speeds up the lookup when forwarding data that is process switched. It also reduces the scope of the queries sent out by a router. If a router has no feasible successor, it queries its neighbor for an alternative route. If the neighbor has no route to offer, the query is forwarded on until a route is found or the search is exhausted. If summarization has been configured, the route that is being queried might have been summarized, and thus the query will end.

14. *Why would you configure an EIGRP router as a stub router?*

The stub configuration is typically used on small-capacity routers in a hub-and-spoke WAN environment. The router has no other neighbors and accesses the network through a distribution layer router. It is not necessary, therefore, for this remote router to have a complete routing table that might overwhelm its limited resources. The remote router needs only a default route to the distribution router that can serve all its needs.

Another reason to configure the remote router as a stub is to lend a hand to the rest of the network. If a query is sent to a remote router, the delays involved might result in the path being SIA. If the stub configuration has been configured, the router will immediately respond to queries as inaccessible.

15. Explain the parameters **receive-only** , **connected** , **static**, and **summary** used in the command **eigrp stub** .

Descriptions of the parameters are listed in the following table.

```
router(config-router)# eigrp stub [receive-only | connected | static | summary]
```

Parameter	Description
receive-only	(Optional) Sets the router as a receive-only neighbor
connected	(Optional) Advertises connected routes
static	(Optional) Advertises static routes
summary	(Optional) Advertises summary routes

16. When configuring the **variance** command, which routes can be used?

Any feasible successor in the topology table can be used in the variance command.

17. Give two reasons why you might wish to change the Hello timer.

The command **ip hello-interval timer** is typically used to decrease the time between Hellos to ensure that the network is more stable and converges more quickly. Although this increases the amount of bandwidth consumed, it is a minimal cost. It becomes very useful in the WANs, particularly when NBMA clouds are used. EIGRP treats both Frame Relay and SMDS as NBMA technologies, resulting in Hello timers that assume a low bandwidth medium (less than T1 speeds) and that set the timer to 60 seconds by default.

18. What should be considered in terms of EIGRP configuration when there are many VCs entering a WAN and no subinterfaces have been configured?

If the serial line has many VCs in a multipoint configuration, EIGRP will evenly distribute its overhead between the VCs, without the use of subinterfaces. The **bandwidth** command should therefore reflect the access link speed into the Frame Relay cloud. If the serial interface is accessing an NBMA environment such as Frame Relay, the situation is straightforward. Your company might have five VCs from your router’s serial interface, each carrying 56 kbps. The access link will need a capacity of $5 * 56$ kbps. Remember, the aggregate configured bandwidth cannot exceed the access speed of the interface.

If the multipoint network has differing speeds allocated to the VCs, a more complex solution is needed. You should take the lowest CIR and to simply multiply it by the number of circuits. This is applied to the physical interface. The problem with this configuration is that the higher-bandwidth links will be underutilized for some things.

Another solution is to configure many subinterfaces, each with a point-to-point link.

19. *On what occasions should you consider configuring the bandwidth on subinterfaces?*

If the multipoint network has differing speeds allocated to the VCs, it is much easier to configure and manage an environment that has used subinterfaces, where a VC is logically treated as if it were a separate interface or point-to-point. The **bandwidth** command can be configured on each subinterface, which allows different speeds on each VC. In this second solution, subinterfaces are configured for the links with the differing CIRs. The links that have the same configured CIR are presented as a single subinterface with a bandwidth, which reflects the aggregate CIR of all the circuits. Cisco recommends this as the preferred solution.

Chapter 15

“Do I Know This Already?” Quiz

1. *b*
2. *a, d*
3. *c*
4. *d*
5. *a*
6. *a, b, d*
7. *a*
8. *c*
9. *d*
10. *b*
11. *d*
12. *a*
13. *b, c*
14. *a*

Q&A

1. *If the weight attribute is used, is a higher or lower weight preferred?*

The weight attribute is proprietary to Cisco and is determined locally on the router. The preference in selection is to the highest weight on the router.

2. *What is an alternative to using BGP as the method of connection to the ISP?*

The alternative method, suggested by Cisco, is to use a default route into the ISP and for the ISP to configure static routes into your autonomous system.

3. *What does the command **clear ip bgp *** achieve, and why should it be used cautiously?*

The command resets BGP peer associations, clears the BGP routing table, and reestablishes BGP connections to the neighbors. It should be used cautiously because the loss of connections will drop packets.

4. *Give three situations in which you should not use BGP to connect to the Internet.*

It is ill-advised to use BGP in certain conditions:

- When the company has only one connection into the Internet
- When there are limited resources on the network
- When the user is not familiar with BGP configurations or policy-based routing
- When the routers do not understand or need to have greater understanding of each other

5. *Explain the use of the command **neighbor 10.10.10.10 remote-as 250**.*

The command tells the router the IP address of the BGP neighbor and the autonomous system to which the neighbor belongs. This information allows the router to create a TCP session with the neighboring router and exchange BGP routing information.

6. *Explain briefly the purpose of the community attribute.*

The community attribute is used to identify prefixes that have some common similarity, without regard to geographic location.

7. *In the route selection process, place the following in order of preference: origin code, weight, local preference, and MED. State the method of selection for the individual attributes themselves.*

The correct order is: highest weight, highest local preference, lowest origin code, and lowest MED.

8. *What is a mandatory attribute?*

A mandatory attribute is an attribute that is well known. It contains information required in BGP messages in order to maintain the BGP network. These attributes are required and therefore recognized by all BGP implementations.

9. *What type of routing protocol is BGP classified as, and what does this mean?*

BGP is classified as an External Gateway Protocol (EGP), as opposed to OSPF, EIGRP, RIP, and so on, which are known as Interior Gateway Protocols (IGPs). If required, BGP can send a summary of the networks known within an organization to maximize security and minimize bandwidth overhead. It is used to convey routing information between autonomous systems.

10. *Explain how static routes can be used as an alternative to configuring BGP to connect to another autonomous system.*

A static route could be configured in conjunction with a default route pointing in the other direction. This is a simple solution when there is not a great deal of information to be exchanged between the autonomous systems. A common example is an organization connecting to an ISP.

A static route is a route that has been manually configured. It has the lowest administrative distance of either 0 or 1 by default, depending on the configuration. This means that it will always take precedence over dynamically learned routes and that it must be redistributed into a routing protocol for other routers to make use of it.

11. *Explain how a default route can be used as an alternative to configuring BGP to connect to another autonomous system.*

A default route is a route used when there is no entry for the remote network in the routing table. It is used to connect to the Internet and other routing domains when it is not practical to know all the available networks. It is sufficient to have an exit point from your network identified.

12. *State two attributes of BGP.*

The following are BGP attributes: next hop, AS_Path, local preference, Multiple Exit Discriminator (MED), community, atomic aggregate, aggregator, and origin.

13. *State four message types of BGP.*

The four message types of BGP are:

- Update messages: Contain paths to destination networks and their attributes. Routes that are no longer available or withdrawn routes are included in updates.
- Open messages: Used to establish connections with peers.
- Keepalives: Sent periodically between peers to maintain connections and verify paths held by the router sending the keepalive.
- Notification: Used to inform the receiving router of errors.

14. *What is policy-based routing?*

Policy-based routing is the means by which traffic can be forced to take a different route from that determined by the dynamic routing protocol. It is defined on a hop-by-hop basis in that the policy is stated on a router and determines which next hop will be used. The decisions can be based on source and/or destination.

15. *What do the letters MED represent? Give a brief explanation of what MED does.*

The Multiple Exit Discriminator is an optional, nontransitive BGP attribute. It is sent only to external BGP peers and is used to influence routers in another autonomous system on the path to take into the autonomous system if multiple paths are available. The lower the value of the attribute, the higher the likelihood that the path will be chosen. By default, a router compares only the MED from routers that are in the same autonomous system as each other but in a different autonomous system from the determining router.

16. *What is a community in BGP?*

A community is a group of networks that share a common property. The commonality is defined by the optional transitive attribute, and it has no physical boundaries. A network can be a member of more than one community. When the community is defined, decisions or filtering can be made based on the group instead of the individual.

17. *Give two reasons why BGP peer groups are useful.*

A peer group is a group of BGP neighbors that share characteristics. The use of peer groups simplifies the configuration of BGP because one configuration effectively configures every router in the peer group. They are also more efficient because updates are generated once per peer group instead of on a per-router (peer) basis, reducing the resources required to support BGP.

18. *What is the difference between a peer and a neighbor?*

In external BGP, there is no difference between a neighbor and a peer. A peer is the BGP term for a neighbor. Both terms refer to a router that is directly connected, with which routing information is exchanged. In iBGP, these routers are not necessarily physically adjacent, but they are the next logical hop router running BGP.

19. *In BGP, describe the purpose of the **network** command.*

The **network** command permits BGP to advertise a network if it is present in the routing table. It is not responsible for starting the BGP process on an interface; instead, it identifies which networks the router originates.

20. Explain the command **neighbor** *{ip-address | peer-group-name}* **next-hop-self**.

On a broadcast multiaccess network such as Ethernet, the next-hop address is the IP address of the advertising router. This command forces BGP to advertise itself as the next-hop router instead of letting the protocol determine the address to be used. This avoids problems seen on NBMA networks or nonmeshed environments, such as X.25 and Frame Relay.

Chapter 16

“Do I Know This Already?” Quiz

1. *d*
2. *a*
3. *d*
4. *a*
5. *d*
6. *b*
7. *c*
8. *a, c, d*
9. *b*
10. *b, c, d*
11. *a*
12. *b*
13. *a, c, d*
14. *a, b, c*
15. *b*
16. *b*
17. *b, c, d*
18. *d*

Q&A

1. *If a route reflector hears an update from a nonclient, what action will be taken?*

If a route reflector hears an update from a nonclient, it will reflect the update to clients only.

2. *In version 11.0 of the Cisco IOS software, what method would be used to restrict routing information from being received or propagated?*

Distribute lists would be used to restrict BGP routing updates in version 11.0 of the Cisco IOS software. Prefix lists became available to ISPs in 11.2 of the IOS and in 12.0 to the general public.

3. *Explain the purpose and use of the command **show ip prefix-list** name [seq seq-number].*

This command displays the entry prefix list with a given sequence number.

4. *Why would you redistribute static routes into BGP?*

The static routes are forwarded to the nonexistent interface null 0 to aggregate the routes to create a supernet. This is then redistributed to the BGP process instead of the hundreds of routes that exist. Because any route redistributed into BGP must be in the IP routing table, this is a way of creating a supernet and having it redistributed.

A static route to null 0 is not needed if the network command is used and no address aggregation is performed, although it can be used to prevent routing loops.

5. *Why is it advisable to have the route reflectors fully meshed?*

Although clients are not fully meshed within a cluster, it is important that the route reflectors they serve are fully meshed. This is to ensure that the routers pass routing information to each other. Remember that iBGP does not propagate routing information; it generates information only based on the network command. The route reflectors need to pass information to each other to ensure the entire network has a full knowledge of the network. It is possible to create a hierarchical design with route reflectors, but this should be done with great caution, as it can cause routing loops.

6. *Why is filtering often required when redistributing BGP into an IGP?*

Filtering is often required when redistributing BGP into an IGP because the routing tables can become overwhelmed by the number of routes that are imported.

7. *What are the advantages of multihoming?*

Multihoming allows for redundancy, which is important when the link is into the Internet and can carry crucial business information for the company. Multihoming also increases performance by allowing the selection of the better paths.

8. *Why do iBGP peers need to be fully meshed?*

iBGP needs to be fully meshed because it has a fundamental rule stating that iBGP will not propagate back into iBGP a route that it has learned from iBGP. The reason is that this is the only way to prevent routing loops. If each router is connected to every other router (fully meshed), the updates come directly from the source, which removes the need for a router to propagate any updates that it has received.

9. *How is a fully meshed network avoided in iBGP?*

A fully meshed network can cause some problems to the routers because they have a great deal of information to handle from every directly connected router. The configuration of route reflectors does not use as much bandwidth, CPU, or memory. Route reflectors allow the router to pass routes on to its peers, and a hub-and-spoke topology is thus possible.

10. *What is the equation to determine the number of sessions needed in a fully meshed BGP network?*

The equation for determining the number of sessions required is $n(n-1)/2$, where n is the number of routers. Thus, 10 routers would mean $10(10-1)/2 = 10 * 9 / 2 = 45$ sessions.

11. *Why does a fully meshed network in iBGP cause problems?*

A fully meshed network in iBGP causes problems because the network has to accommodate a large number of TCP connections, and this can eat up memory, CPU, and eventually bandwidth.

12. *State two benefits to using route reflectors.*

Route reflectors have many benefits:

- The use of a router reflector means that fewer TCP peer connections are needed. This streamlines the network traffic and solves the excessive use of network resources sometimes incurred with a fully meshed network.
- The design and configuration are very straightforward, which means that it is easy to implement and thus to migrate an existing network, particularly because path attributes are not affected by them.
- Despite the fact that the route reflectors are straightforward, the flexibility in the design means that it is possible to become very sophisticated using redundant route reflectors and even multiple levels of route reflectors. Complex solutions are possible using route reflectors.

- 13.** *If a route reflector sees multiple paths to a destination, what action is taken?*

The route reflector chooses the best path to the destination.

- 14.** *Explain the difference between a cluster-ID and an originator-ID.*

The cluster-ID is an optional, nontransitive BGP attribute (type code 10). A cluster is the route reflector and its clients. If a cluster has more than one route reflector, it needs to be identified by a cluster-ID. The cluster-ID is used to identify the route reflector advertising the prefix. This is listed in the update; if a router receives an update with its own cluster-ID, it will be ignored, thus preventing routing loops.

An originator-ID is the ID given to the source route-reflector client. It serves the same purpose as the cluster-ID in that it prevents routing loops. A router that receives an update that contains its own originator-ID will ignore the update.

- 15.** *State two advantages in using prefix lists over access lists.*

The advantages of using prefix lists instead of access lists include the following:

- They have a better performance than access lists.
- They allow editing of the lists so that additional lines of code can be inserted anywhere in the list.
- The user interface is easier to use.
- They are more flexible.

- 16.** *If the ISP has provided a default route, how will the router within the autonomous system select the exit path in a multihomed environment?*

The IGP within the autonomous system will use the metric associated with that routing protocol. The router running EIGRP, for example, will select the nearest router based on the composite metric of bandwidth and delay (by default).

- 17.** *What is a disadvantage of an autonomous system receiving full routing updates from all ISPs?*

The disadvantage is that a great deal of network resources, such as memory, bandwidth, and CPU, is required.

18. *What is the danger of redistributing BGP into the IGP?*

The danger is that the autonomous system routers receiving the BGP updates will be overwhelmed by the amount of routing information that they receive. This could result in an unstable network or even a network exhausted to death.

19. *What are the advantages of a fully meshed iBGP network?*

The advantage of a fully meshed network is that the network will receive full routing information from the directly connected peers. This means that the IGP does not need to carry routing information to the BGP peers, and no redistribution is necessary. This does not run the risk of overloading the IGP. It also means that the network will converge more quickly and that synchronization can be turned off, which will improve performance.

20. *In configuring a route reflector, how is the client configured?*

The client of a route reflector is configured at the same time as the route reflector. The command issued on the router that is to become the route reflector identifies the router that is to become the client. The following command is the syntax issued at the router reflector:

```
neighbor ip-address route-reflector-client
```

21. *What commands are used to display the BGP router ID that identifies the router that is sending the updates and peering with its neighbor?*

The commands that display the BGP router ID are `show ip bgp neighbor` or `show ip bgp .`

Chapter 17

“Do I Know This Already?” Quiz

1. *a, c*
2. *a, c*
3. *b*
4. *a*
5. *a, c*
6. *c*
7. *a, b*
8. *c, d*
9. *b*

- 10. *c*
- 11. *b*
- 12. *a, b, c*
- 13. *c*
- 14. *a, b, c*

Q&A

1. *State two of the methods that Cisco recommends for controlling routing protocol traffic.*

The methods that Cisco discusses as useful methods of controlling routing updates are as follows:

- Passive interfaces: Prevent routing protocol communication and thus an adjacency from being formed with neighboring routers.
- Changing the administrative distance on the route: Changes the natural order as laid down by Cisco on which routing protocol is more plausible than the others. A scale of weighting is applied to the protocols.
- Default routes: Instruct the router on where to send the destination traffic if the routing table has no entry for that destination.
- Static routes: Offers the ability to configure manually the path to a destination network.
- Route update filtering: Offers the use of access lists to control the exchange of routing information.

2. *What is the default administrative distance for RIP?*

The administrative distance for RIP is 120; it has the highest distance of interior routing protocols and is therefore the least likely to be selected.

3. *State two instances when you do not want routing information propagated.*

The two occasions that you do not want routing information to be propagated are as follows:

- If there is a WAN link where the cost of the link is based on network traffic. This might also have the added disincentive of being a WAN link that is a dial-on-demand link, which is raised and maintained by the presence of traffic attempting to flow across the interface.

- When you are trying to prevent routing loops. If the routing domain has redundant paths that will be learned by different routing protocol, it is advisable to filter the propagation of one of the paths.

4. *In what instances will EIGRP automatically redistribute?*

EIGRP will automatically redistribute between itself and IGRP as long as both processes are running the same autonomous system number.

EIGRP for IPX automatically redistributes; IPX RIP/SAP and Enhanced IGRP for AppleTalk similarly redistribute automatically into RTMP.

5. *Which command is used to view the administrative distance of a route in the routing table?*

The command **show ip route** displays the administrative distance for each route; two numbers are shown in brackets: the route metric, followed by the administrative distance.

6. *When is redistribution required?*

Redistribution is required when there is more than one routing protocol for IP running within the organization and when every part of the network needs connectivity to all the networks.

7. *Why does Cisco recommend that you not overlap routing protocols?*

The reason not to overlap routing protocols is that it will increase network traffic, router CPU processing, and memory because of the additional protocol updates. This additional traffic and CPU and memory requirements complicate the routing process. The decision is not simply between multiple paths, but also between the various routing protocols that are advertising them.

8. *Why would you want to prevent routing updates across an on-demand WAN link?*

A WAN link that is a dial-on-demand link is raised and maintained by the presence of interesting traffic attempting to flow across the interface. Thus, every time a RIP update is sent, the path would be raised or, more likely, kept up all the time.

9. *What is the metric used for in a routing protocol?*

The metric is used to select the best path when multiple paths are available to a remote network.

10. *Give two reasons for using multiple routing protocols.*

The main reasons for multiple protocols existing within an organization are as follows:

- The organization is transitioning from one routing protocol to another because the network has grown and there is a need for a more sophisticated protocol that will scale.

- Although a vendor solution is preferred, there is a mix of different vendors within the network, so the vendor solution is used in the areas available. This is particularly true in client/server networks.
- Historically, the organization was a series of small network domains that have recently been tied together to form one large enterprise network. The company may well have plans to transition to a single routing protocol in the future.
- Often after a merger or a takeover, when several companies become one, it takes planning, strategy, and careful analysis to determine the best overall design for the network.
- Politically, there are ideological differences among the different network administrators, which until now have not been resolved.

11. *In a very large environment, the various domains might have different requirements, making a single solution inefficient. A clear example is the case of a large multinational corporation, where EIGRP is the protocol used at the access and distribution layer, but BGP is the protocol connecting the core. When implementing redistribution, state one possible problem that you might experience, and explain why it is a problem.*

The problems experienced as a result of multiple routing processes and their redistribution include the following:

- The wrong, or less efficient, routing decision is made, referred to as the suboptimal path.
- A routing loop occurs, in which the data traffic is sent in a circle without ever arriving at the destination.
- The convergence time of the network increases because of the different technologies involved. If the routing protocols converge at different rates, this can also cause problems. In some cases, this might result in timeouts and the temporary loss of networks.

12. *Which has a lower administrative distance, IGRP or OSPF?*

IGRP has an administrative distance of 100, whereas OSPF has an administrative distance of 110. The IGRP path will be entered into the routing table if there are paths offered to the same destination from both protocols.

13. *What command is used to configure an outbound route filter?*

The command for configuring a route filter is as follows:

```
Router(config-router)#distribute-list {access-list-number | name} out
[interface-name | routing-process | autonomous-system-number]
```


14. *What is a passive interface?*

A passive interface is an interface that will listen to routing updates but that will not propagate any updates for the protocol configured, although OSPF and EIGRP will not learn routes over a passive interface. It is used to prevent unnecessary traffic from being sent out of an interface. Usually a passive interface is configured when there are no routers to hear the updates on that network.

15. *What is the purpose of administrative distance?*

When the routing table is populated with networks that are provided by multiple routing protocols, the administrative distance is used to choose the best path to the remote network.

16. *What is the concern of redistributing into a redundant network?*

The concern with redistributing into a redundant network is that it is possible to select suboptimal routes. That is, if two routing protocols both have a path to the same destination, the path for the routing table will be selected based on administrative distance. This is a blanket solution that does not always render the best decision.

17. *What is a default network?*

A default network is a route used when there is no entry for the remote network in the routing table. It is used to connect to the Internet and other routing domains when it is not practical to know all the available networks. It is sufficient to have an exit point from your network identified.

18. *Why is it necessary to configure a default metric when redistributing between routing protocols?*

The metric is used within a routing protocol to select the best path to a remote network when there are multiple paths. When redistributing, it is not always possible to port the metric across because the metric is protocol-specific. The default metric throws the original metric away and substitutes a new metric for the new routing protocol.

19. *Which command is used to modify the administrative distance of a route?*

The command to configure the administrative distance of a route is as follows:

```
Router(config-router)#distance weight [address mask] [access-list-number |
name] [ip]
```

A different command is used for EIGRP and BGP. The EIGRP command to change the administrative distance is as follows:

```
Router(config)#distance eigrp internal-distance external-distance
```

20. *What is the difference in processing for an inbound and an outbound route filter?*

If a filter is applied to an incoming interface, the routing table is not checked unless the route is permitted. However, the router has to examine every incoming packet, which is expensive in resources. The outbound filter must go through the routing decision process the first time, after which the result is cached.

Inbound filters are wider ranging because they prevent routes from entering the router instead of filtering on each outgoing interface that is affected.

Chapter 18

“Do I Know This Already?” Quiz

1. *c*
2. *a, b, d*
3. *b*
4. *d*
5. *a*
6. *a, c, d*
7. *c*
8. *b*
9. *d*
10. *a*
11. *a*
12. *c*
13. *c*
14. *b*
15. *c*
16. *a*
17. *b*

Q&A

1. *Explain the command **match ip address** {access-list-number | name} [access-list number | name].*

The command is used to match criteria in establishing the policy-based routing. Access lists are used to specify the addressing of the packets to be affected.

2. *Explain the command **ip route-cache policy** .*

This command is set on an incoming interface and enables the fast switching of policy-based routing. Before version 11.2 of the Cisco IOS software, policy-based routing was process-switched. This caused some applications to time out, but the problem has now been resolved. Fast switching of policy-based routing is disabled by default. Therefore, it is necessary to manually configure it.

3. *State two benefits of using policy-based routing.*

The benefits of policy-based routing include the following:

- Organizations can determine traffic flow based on the origin of the traffic. They can send traffic owned by different groups across different paths.
- QoS can be set in the IP header using the precedence or TOS bits. This allows certain traffic to be prioritized through the network.
- High-cost links can be raised or made active on more specific criteria, which allows an efficient use of the resources available.
- Traffic can be sent across multiple paths based on traffic characteristics.

4. *How are matching routes modified in a route map?*

Using the **set** command modifies matching routes. If the criteria are met in the **match** command and the action was to permit, the set criteria is initiated to control the routing as specified.

5. *Explain the command **set ip default next-hop** [ip-address...ip-address].*

This command provides a list of IP addresses for traffic if there is no explicit route in the routing table for the destination address of the packet. These addresses are those of next-hop routers or of the interfaces of adjacent routers. If multiple next-hop addresses are listed, then the first address is tried. If it is unavailable, the others are tried in turn.

6. *Which command displays route maps that are configured on interfaces?*

The command **show ip policy** displays the route maps used for policy-based routing on the router's interfaces. The command **show route-map** [map-name] displays the route maps.

7. *What command is used to attach a route map to an incoming interface?*

The command **ip policy route-map map-tag** assigns the route map to the incoming interface where it examines all incoming packets and issues set commands on all packets that match at least one of the match criteria.

8. *What is a map tag?*

A map tag is the name of the route map to be used in policy-based routing. This must match a map tag specified by a route-map command.

9. *Explain briefly the difference between the **match** and **set** commands.*

The **match** command determines whether the packet will be routed using the route map, and the **set** command determines how the packet will be routed.

10. *What are the criteria by which policy-based routes are determined?*

Instead of routing by the destination address, policy-based routing allows you to determine and implement routing policies to allow or deny paths based on the following:

- The identity of a particular end system
- The application being run
- The IP protocol in use
- The size of packets

11. *How would you block traffic that found no match?*

The way to block traffic that is not matched in the route map is to add an extra line of configuration. Adding a set command at the end of the route map routes all unmatched routes to the interface null0.

12. *What is the purpose of the sequence number in a route map?*

Within a route map, each route map statement is numbered with sequence numbers and, therefore, can be edited individually. The sequence number is also used to specify the order in which conditions are checked. Thus, if there are two statements in a route map named BESTEST, one with sequence 5 and the other with sequence 15, sequence 5 is checked first. If there is no match for the conditions in sequence 5, then sequence 15 will be checked.

13. *What logic is used if there are multiple match statements in the route map?*

A route map statement might contain multiple match statements. All match statements in the route map statement must be considered true for the route map statement to be considered matched. This is a logical AND.

14. *What parameters can an extended access list define in the route map selection process?*

An extended IP access list is used to specify criteria based on source and destination, application, protocol type, TOS, and precedence. If multiple access list statements are specified, matching any one will result in a match.

15. *How can policy-based routing save money for the network?*

The very nature of policy-based routing allows the network to be streamlined and to forward traffic down cheaper circuits. The bulk traffic generated by a specific activity can be diverted to use a higher-bandwidth, high-cost link for a short time. Meanwhile, interactive traffic is provided basic connectivity over a lower-bandwidth, low-cost link. For example, a dial-on-demand ISDN line might be raised in response to traffic to a finance server for file transfers selected by policy-based routing.

16. *What are some of the potential disadvantages of using policy-based routing?*

The following are disadvantages of policy-based routing:

- You need a backup path in place in case the defined next-hop router goes down. If there is no alternative defined, policy-based routing will default to dynamic routing decisions.
- Additional CPU is required to examine every source address to affect the defined policy.
- Extra configuration is required.
- The possibility exists that other traffic will be disrupted.

17. *Explain the use of the command **match length min max**.*

This command is used to define the criteria based on the Layer 3 length of the packet. The *min* parameter states the minimum inclusive length of the packet allowed for a match. The *max* parameter states the maximum inclusive length of the packet allowed for a match.

In this way, interactive traffic that is time-sensitive, such as SNA traffic tunneled in IP, can be sent on a dedicated route. Interactive traffic uses small packets, so the links could be dedicated

by packet size, allowing file transfers using large packets to use a separate link so that the terminal sessions are not starved of resources.

18. *Explain the following sample configuration.*

```
Router(config)#interface s0
Router(config-if)#ip address 190.10.10.2 255.255.255.0
Router(config)#ip policy route-map tryout
Router(config)#!
Router(config)#access-list 5 permit 140.10.0.0 0.0.255.255
Router(config)#!
Router(config)#route-map tryout permit 10
Router(config-route-map)#match ip address 5
Router(config-route-map)#set interface e1
Router(config-route-map)#set ip default next-hop 190.10.20.2
```

The serial interface has been configured with an IP address and has the route map tryout assigned to make decisions about inbound packets.

The route map tryout is configured to send all packets that are destined for network with the address in the 140.10.0.0 network address space. All routes that have this destination address will be sent out of the interface e1, whereas all unknown addresses, including broadcasts, will be sent to the default next hop of 190.10.20.2.

19. *Briefly explain the use of the command **show ip policy** .*

This command displays the route maps used for policy-based routing on the router’s interfaces.

20. *Configuring route maps is complex, and it is easy to confuse the logic by which they work. State one of the things you should be aware of when configuring a route map.*

When editing a route map statement with the **no** version of the existing command line, if you forget to type in the sequence number, you will delete the entire route map.



Suggested Reading

This appendix provides a list of supplemental and background reading. All the references listed here are excellent resources. The nature of the information is diverse; some of the books, for example, are tomes of wisdom, good reference guides about command syntax, or a consideration of a very specialized problem that you might encounter.

The list is broken up into general topic areas and further divided by information type—books, web sites, and, where possible, standards documents.

NOTE If a web page is not accessible, the specific location might have changed. If this happens, use a search engine to hunt down the new location. Many of the web references are to the Cisco site, so finding an updated URL, if necessary, should not prove too problematic.

General TCP/IP

Books

Comer, Douglas E. *Internetworking with TCP/IP*, Volume 1. Prentice Hall; 2000.

Stevens, W. Richard. *TCP/IP Illustrated: The Protocols*, Volume 1. Addison-Wesley; 1994.

Doyle, Jeff. *Routing TCP/IP*, Volume I (CCIE Professional Development). Cisco Press; 1998.

Standards Documents

RFC 3056, “Connection of IPv6 Domains via IPv4 Clouds.” B. Carpenter and K. Moore.

RFC 2893, “Transition Mechanisms for IPv6 Hosts and Routers.” R. Gilligan and E. Nordmark.

RFC 2373, “IP Version 6 Addressing Architecture.” R. Hinden and S. Deering.

RFC 2365, “Administratively Scoped IP Multicast.” D. Meyer.

Web Sites

“IPv6 Deployment Strategies” at http://www.cisco.com/en/US/tech/tk872/tk373/technologies_white_paper09186a00800c9907.shtml

“ABCs of IP version 6” at <http://www.cisco.com/warp/public/732/abc/docs/abcipv6.pdf>

IP Routing Protocols

Books

Huitema, Christian. *Routing in the Internet*. Prentice Hall PTR; 2000.

Perlman, Radia. *Interconnections: Bridges, Routers, Switches, and Internetworking Protocols*, Second Edition. Addison-Wesley; 1999.

Shamim, Faraz, Zaheer Aziz, Johnson Liu, and Abe Martey. *Troubleshooting IP Routing Protocols* (CCIE Professional Development). Cisco Press; 2002.

Web Sites

“Designing Large-Scale IP Internetworks” at <http://www.cisco.com/univercd/cc/td/doc/cisintwk/idg4/nd2003.htm>

“IP & Routing” at <http://www.cisco.com/warp/public/732/Tech/routing/>

OSPF

Books

Moy, John. *OSPF: Anatomy of a Routing Protocol*. Addison-Wesley; 1998.

Standards Documents

RFC 2370, “The OSPF Opaque LSA Option.” R. Coltun.

RFC 2328, “OSPF version 2.” J. Moy.

RFC 1586, “Guidelines for Running OSPF Over Frame Relay Networks.” O. deSouza and M. Rodrigues.

Web Sites

“Problems with Running OSPF in NBMA Mode over Frame Relay” at <http://www.cisco.com/warp/public/104/24.html>

“Configuring OSPF” chapter of the Network Protocols Configuration Guide, Part 1, at http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/12cgcr/np1_c/1cprt1/1cospf.htm

“OSPF Commands” at http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/12cgcr/np1_r/1rprt1/1rospf.htm

“Open Shortest Path First” at http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/ospf.htm

IS-IS

Books

Martey, Abe. *IS-IS Network Design Solutions*. Cisco Press; 2002.

Standards Documents

RFC 1195, “Use of OSI IS-IS for Routing TCP/IP and Dual Environments.” R. Callon. December 1990.

RFC 1629, “Guidelines for OSI NSAP Allocation in the Internet” (Obsolete RFC 1237.) R. Colella, R. Callon, E. Gardner, Y. Rekhter. May 1994.

RFC 1142, “OSI IS-IS Intra-domain Routing Protocol.” D. Oran. February 1990.

ISO 10589, “Intermediate System to Intermediate System Intra-Domain Routeing Exchange Protocol for use in Conjunction with the Protocol for Providing the Connectionless-mode Network Service (ISO 8473).” April 1992.

Web Sites

“Integrated IS-IS Commands” at <http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122sup/122csum/csum1/122csip2/1sfisis.htm#xtocid19>

EIGRP

Books

Retana, Alvaro, Don Slice, and Russ White. *Advanced IP Network Design* (CCIE Professional Development). Cisco Press; 1999.

Web Sites

“EIGRP—Enhanced Interior Gateway Routing Protocol” at http://www.cisco.com/en/US/tech/tk365/tk207/technologies_white_paper09186a0080094cb7.shtml

“Configuring IP Enhanced IGRP” at http://www.cisco.com/en/US/products/sw/iosswrel/ps1835/products_configuration_guide_chapter09186a00800ca762.html

BGP

Books

Halabi, Sam, and Danny McPherson. *Internet Routing Architectures*, Second Edition. Cisco Press; 2000.

Standards Documents

RFC 1774, “BGP-4 Protocol Analysis.” P. Traina. March 1995.

RFC 1771, “A Border Gateway Protocol 4 (BGP-4).” Y. Rekhter and T. Li. March 1995.

RFC 1163, “A Border Gateway Protocol (BGP).” K. Lougheed and Y. Rekhter. June 1990.

Web Sites

“Border Gateway Protocol (BGP)” at http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm

“Configuring BGP” at http://www.cisco.com/univercd/cc/td/doc/product/software/ios121/121cgcr/ip_c/ipcprt2/1cdbgp.htm

“BGP Case Studies” at <http://www.cisco.com/warp/public/459/bgp-toc.html>

Redistribution

Books

Benjamin, Henry. *CCNP Practical Studies: Routing*. Cisco Press; 2002.

Web Sites

“Configuring IP Routing Protocols” at http://www.cisco.com/en/US/products/sw/iosswrel/ps1818/products_configuration_guide_chapter09186a008008795f.html

“RIP and OSPF Redistribution” at <http://www.cisco.com/univercd/cc/td/doc/cisintwk/ics/cs001.htm>



Glossary

ACK An acknowledgment of packets sent reliably.

Active EIGRP state when a network change is seen, but on interrogation of the topology table, there is no feasible successor. The router queries its neighbors for alternative routes.

adjacency State when a router understands a partner with whom it can exchange protocol traffic. In link-state routing protocols, adjacencies are formed when two neighboring routers have exchanged Hellos. Each routing protocol has different rules about adjacencies. For example, in EIGRP, an adjacency is formed on receipt of a neighbor's Hello. In IS-IS, it is local routing information that shows the reachability of a directly connected end system (ES) or intermediate system (IS). A separate adjacency is created for each neighbor on a circuit and for each level of routing (that is, Level 1 and Level 2) on a broadcast circuit.

adjacent neighbors A neighbor is a router that is directly connected to another router. They must also have same mask and Hello parameters on the connecting link. An adjacent router is a router that has exchanged routing information with its neighbor.

administrative distance A rating of the trustworthiness of a routing information source. The higher the value, the lower the rating of trust; it is an inverse rating. This is used when there are multiple routing protocols updating the IP routing table and when a method other than metric is required to select the best route.

Administrative Domain A group of routers that share the same routing protocol within one organization. Also known as an *autonomous system*.

advertised distance (AD) In EIGRP, it is the cost of the path to the remote network from the neighbor (the metric from the next-hop router).

aggregated route The consolidation of advertised addresses in a routing table. Aggregating routes reduces the number of routes in the routing table, the routing update traffic, and overall router overhead. This is also called "route summarization."

area A logical set of network segments and their attached devices. Areas are usually connected to other areas by means of routers, making up a single autonomous system. Used in DECnet,

IS-IS, and OSPF, it is a subdomain within an Administrative Domain. Routers in an area maintain detailed routing information about the area's internal composition. The routers also maintain routing information that allows them to reach other areas. In IS-IS, the area address is contained in the NET and NSAP address.

Area Border Router (ABR) In OSPF, this router is responsible for connecting two or more areas. It holds a full topological database for each area to which it is connected and sends LSA updates between the areas. These LSA updates are summary updates of the subnets within an area. A similar function of connecting areas is provided by the Level 1-2 router in IS-IS.

attribute BGP parameters used to select the best path. In essence, the attributes are the metrics used by BGP. Examples of attributes are local-preference, weight, and autonomous system-path.

autonomous switching Feature on Cisco routers that provides faster packet processing by allowing the ciscoBus to switch packets independently without interrupting the system processor.

autonomous system Definition for the organizational boundary. A collection of networks under a common administration sharing a common routing strategy. Autonomous systems can be subdivided into areas. Within the terminology of the routing protocols, it defines all the routers within an administrative domain, where each router has full knowledge of the subnets within the domain. If you are connecting directly to the Internet using BGP, the autonomous system number must be unique and obtained from the Internet addressing committees.

backup designated router (BDR) In OSPF, this is the backup to the designated router (DR), in case the DR fails. The BDR performs none of the DR functions while the DR is operating correctly.

Berkeley Software Distribution (BSD) Term used to describe any of a variety of UNIX-type operating systems based on the UC Berkeley BSD operating system.

circuit In IS-IS, this is the local routing information for a single subnet point of attachment (SNPA).

Cisco Express Forwarding (CEF) Advanced, Layer 3 IP switching technology. CEF optimizes network performance and scalability for networks with large and dynamic traffic patterns, such as the Internet, on networks characterized by intensive Web-based applications or interactive sessions.

classful routing protocols Routing protocols that do not transmit any information about the prefix length. A distance vector routing protocol that will not allow VLSM or route summarization. Examples are RIP and IGRP.

classless interdomain routing (CIDR) This is the means by which the Internet assigns blocks of addresses, typically Class C addresses, and summarizes them by using the prefix mask.

classless routing protocols Routing protocols that include the prefix length with routing updates; routers running classless routing protocols do not have to determine the prefix themselves. Classless routing protocols support VLSM.

cluster-ID In BGP, the cluster-ID is another attribute used in configuring route reflectors. If the cluster has more than one route reflector, the cluster-ID is used to recognize updates from other route reflectors within the cluster.

Code/Length/Value (CLV) In IS-IS, these are the variable-length fields in a PDU. The Code field specifies the information in the Content field as a number. The Length field states the size of the Value field. The Value field contains the information itself.

complete sequencenumber packet (CSNP) In IS-IS, CSNPs describe every link in the link-state database. CSNPs are sent on point-to-point links when the link comes up to synchronize the link-state databases. The designated router (DR), or designated intermediate system (DIS), on a multicast network sends out CSNPs every 10 seconds.

Connectionless Network Protocol (CLNP) This is the ISO protocol used to carry data and error indications at the network layer. CLNP is similar to IP and has no facilities to detect errors in data transmission. It relies on the transport layer to provide guaranteed data delivery.

Connectionless Network Service (CLNS) CLNS uses a datagram transfer service and does not require a circuit to be established before data is transmitted.

connection-oriented Software on two end nodes guarantees the transmission of network traffic because a circuit setup is established before sending any data. It requires the use of sequencing, windowing, and acknowledgements.

convergence Speed of a group of internetworking devices running a specific routing protocol to agree on the topology of an internetwork after a change in that topology.

cost The metric for OSPF. It is not defined in the standard with a value. Cisco uses the default of the inverse of bandwidth so that the higher the speed of the link, the lower the cost—and, therefore, the more attractive the path.

count to infinity Problem that can occur in routing algorithms that are slow to converge, in which routers continuously increment the hop count to particular networks. Typically, some arbitrary hop count limit is imposed to limit this problem.

database descriptor In OSPF, the database descriptor is referred to as DBDs or database descriptor packets (DDPs). BSCI literature uses the initialization DBD. These are packets exchanged between neighbors during the exchange state. The DBDs contain LSAs, which describe the links of every router in the neighbor's topology table.

default routes A route that should be used if the destination network is not otherwise present in the routing table.

designated intermediate system (DIS) In IS-IS, the router (IS) on a LAN that is designated to perform additional duties. In particular, the DIS generates link-state PDUs on behalf of the LAN by treating the LAN as a pseudonode.

designated router (DR) In OSPF, this is the router responsible for making adjacencies with all neighbors on a multiaccess network, such as Ethernet or FDDI. The DR represents the multiaccess network, in that it ensures that every router on the link has the same topology database.

Diffusing Update Algorithm (DUAL) In EIGRP, this is the algorithm performed on the topology table to converge the network that provides loop-free operation at every instant throughout a route computation. This allows routers involved in a topology change to synchronize at the same time, while not involving routers that are unaffected by the change.

Dijkstra algorithm Routing algorithm that iterates on length of path to determine a set of shortest paths in a shortest-path tree. Commonly used in link-state routing algorithms. Sometimes called “shortest path first algorithm.” This algorithm is used by OSPF.

distance-vector routing protocol Class of routing algorithms that have neighbors exchange routing tables and use the Bellman-Ford algorithm to find shortest paths. Distance vector routing algorithms can be prone to routing loops but are computationally simpler and require fewer resources than link-state routing algorithms.

distribute list An access list that is applied to the routing protocol. It is used to control routing updates by filtering out those routes that are not to be propagated. This is particularly useful in preventing routing loops in redistributed networks.

Domain Name System (DNS) System used on the Internet for translating names of network nodes into IP addresses.

dot address Refers to the common notation for IP addresses in the form *n.n.n.n*, where each number *n* represents, in decimal, 1 byte of the 4-byte IP address. This is also called “dotted notation” or “four-part dotted notation.”

dotted decimal notation Syntactic representation for a 32-bit integer that consists of four 8-bit numbers written in base 10 with periods (dots) separating them. It is used to represent IP addresses on the Internet, as in 192.67.67.20. This is also called “dotted quad notation.”

Dual IS-IS IS-IS that supports both OSI and IP routing information. Areas within the autonomous system can run OSI, IP, or both. However, the configuration chosen must be consistent within the entire area.

Dynamic Host Configuration Protocol (DHCP) Provides a mechanism for allocating IP addresses dynamically so that addresses can be reused when hosts no longer need them.

dynamic routes Automatic rerouting of traffic based on sensing and analyzing current actual network conditions, not including cases of routing decisions taken on predefined information.

end node A device that is connected to the network.

end system (ES) The end node or host, which has limited routing capabilities. The ES has the OSI or IP Layer 3 protocol running and can receive and send data.

End System-to-Intermediate System (ES-IS) In IS-IS, this is the protocol by which the OSIES and the IS communicate to dynamically learn Layer 2 adjacencies.

exchange state In OSPF, this is the method by which two neighboring routers discover the map of the network. When these routers become adjacent, they must first exchange DBDs to ensure that they have the same topology table.

exstart In OSPF, this is the state in which the neighboring routers determine the sequence number of the DBDs and establish the master/slave relationship.

Exterior Gateway Protocol (EGP) Protocol that runs between autonomous systems. There is also a protocol with this name that was the precursor to BGP.

exterior routing A routing protocol used to exchange information between autonomous systems or organizations, used to connect organizations into the Internet. BGP and EGP are examples of exterior routing protocols.

external BGP (eBGP or eBGP-4) When BGP is used to connect different autonomous systems.

fast switching A cache in the Cisco router that contains routing decisions. After the routing decision for a packet has been made, it can be cached in any one of a variety of caches. This means that the forwarding of traffic through the router is greatly enhanced.

feasible condition (FC) In EIGRP, this is when a neighbor reports a path cost (AD) that is lower than the router's FD to a network.

feasible distance (FD) In EIGRP, this describes the lowest-cost distance (metric) to a remote network.

feasible successor (FS) A term used by EIGRP to describe a next-hop router with a path to the remote network that EIGRP considers a viable route, one guaranteed to be loop-free. The router that is an FS must report an AD lower than the FD of the router, thus meeting the FC.

first octet rule The mechanism by which the Layer 3 device identifies the class of IP address. If the protocol is a classful protocol, the first octet rule is the only means available to determine the network portion of an address to which it is not directly connected.

flash update A routing update sent asynchronously in response to a change in the network topology. If there is a change in the metric, the update is sent immediately without waiting for the update timer to expire. Sometimes known as *triggered updates*.

floating static route A floating static route is a route that, although manually configured, has been identified as a route to choose only if the dynamically learned routes fail. These routes need to have a higher administrative distance than the dynamically learned route so that they can be overridden by dynamically learned routing information. This is used to create a DDR backup to an existing link.

flooding A traffic-passing technique used by switches and bridges in which traffic received on an interface is sent out to all the interfaces of that device, except the interface on which the information was originally received.

Hello Hello packets are used to discover and maintain adjacencies. Messages are used to find and maintain neighbors in the topology table. They are sent periodically and unreliably.

hierarchical design Cisco recommends that a hierarchical design be used when designing large networks. The complex problem of routing in a large network is simplified by breaking the network into a hierarchy of networks, where each level is responsible for its own routing. Cisco uses three levels, which it calls the access, distribution, and core levels.

holdtime Value set in the Hello packet. It determines how long the router waits for Hellos from a neighbor before declaring the neighbor unavailable. This information is held in the neighbor table.

host address This is a subset of the ISO NET address, which includes the domain, area, and system IDs.

IANA Address This is a classful address allocated by the IANA.

incremental update A routing update that is sent only when there is a change in the topology, not periodically when a timer expires.

Init An OSPF state in which a Hello packet has been sent from the router, which is waiting for a reply to establish two-way communication.

Integrated IS-IS Another term for Dual IS-IS. Indicates IS-IS can be used to support routing for two Layer 3 protocols (IP and CLNP) in the same network simultaneously.

Interior Gateway Protocol (IGP) In the past, the term “gateway” was used to define a router. This is a routing protocol that runs within an autonomous system.

interior routing protocol A protocol used to route information between routers within an autonomous system or organization.

intermediate system (IS) A router. The IS is a device capable of directing traffic to remote destinations.

Intermediate System-to-Intermediate System (IS-IS) The OSI routing protocol that learns the location of the networks within the autonomous system so that data can be forwarded to the remote hosts.

internal BGP (iBGP or iBGP-4) When BGP is used to connect routers resident in the same autonomous system.

internal router An OSPF router that has all its interfaces in the same area.

internet Short for *internetwork*. Not to be confused with the Internet. See also *internetwork*.

Internet The Internet is commonly defined as a network of networks. It is a worldwide mesh of hundreds of thousands of networks, owned and operated by millions of people all over the world, all connected to thousands of ISPs. It provides the World Wide Web (WWW), a series of Web pages offering a range of services and information.

Internet Assigned Numbers Authority (IANA) Responsible for address allocation in the Internet.

internetwork A collection of networks interconnected by routers and other devices that functions (generally) as a single network. It is sometimes called an internet, which is not to be confused with the Internet.

IPv4 IP version 4 is the current version of IP that is used extensively throughout the world. It is defined in RFC 791.

IPv6 IP version 6 is a replacement for the current version of IP (version 4). IPv6 includes support for flow ID in the packet header, which can be used to identify flows. IPv6 used to be called “IPng (next generation).”

IS-IS domain A group of routers running the IS-IS protocols for exchanging routing information.

Layer 3 switching Used in the context of VLANs, the mechanism by which a switch will route between VLANs. It also refers to routers, when the routing decision has been made and the result has been cached. The subsequent lookup involves switching (for example, fast switching), but on a Layer 3 decision.

Level 1 (L1) Used in IS-IS to reference a router or a link. These links or routers are internal to the area. The routers receive routing information for their area only and have no knowledge of the other areas' networks. To reach other areas, Level 1 routers maintain a default route to the nearest Level 2 router.

Level 1-2 (L1-2) Used in IS-IS to reference a router or a link that connects areas. This router connects a Level 1 area to the Level 2 backbone. It has a Level 1 routing table to route to ES and IS in its own area by system ID. It maintains a Level 2 prefix table to route to other areas.

Level 2 (L2) Used in IS-IS to reference a link or a router. These routers are connected only to the backbone and provide transit traffic between areas.

link A physical connection to a neighbor. This link is then transmitted to all the other routers in the area via the LSP.

link-state advertisement (LSA) In OSPF, this is a packet describing a router's links and the state of those links. There are different types of LSAs to describe the different types of links. It is a broadcast packet used by link-state protocols that contains information about neighbors and path costs. LSAs are used by the receiving routers to maintain their routing tables.

link-state database In OSPF, this is the raw material for constructing a topology map. It is a copy of every LSA in the area, giving a list of every router, its links, and the state of the links. From this information, it can create a SPF tree of every network and every path to each network.

link-state packet (LSP) In IS-IS, this is a packet that describes a router's links. There are separate LSPs for Level 1 and Level 2 updates.

link-state request (LSR) In IS-IS, this is when the router receives a CSNP complete with LSP, and it compares the LSP against the topological database. If either the LSP entry is not present or the entry is older than the DBD, the router will request further information via an LSR.

link-state routing algorithm A routing algorithm in which each router broadcasts or multicasts information regarding the cost of reaching each of its neighbors to all nodes in the internetwork. Compare with *distance vector routing protocol*.

link-state update (LSU) Update sent in response to the LSR. It is the LSP that was requested.

logical AND The operation where two bits are compared and the result is TRUE only if both output bits are TRUE. The mechanism by which a subnet is derived from an IP host address. The router ANDs the subnet mask, in binary, onto the host address, in binary. The result of the logical AND is the subnet address.

neighbor A router on the same link with which an adjacency is formed and routing information is then exchanged.

neighbor table An EIGRP list of every neighbor, including the IP address, the outgoing interface, the holdtime, smooth round-trip time (SRTT), and uptime, or how long since the neighbor was added to the table. This table is built from information on Hellos received from adjacent routers (neighbors).

NetFlow A Cisco solution that enhances the speed of transmission by caching routing decisions.

Network Address Translation (NAT) Mechanism for reducing the need for globally unique IP addresses. NAT allows an organization with addresses that are not globally unique to connect to the Internet by translating those addresses into globally routable address space.

network congestion A condition in which excessive traffic on the network is the cause of delays and packet loss.

network entity title (NET) In IS-IS, this is the ISO address of the system, but not to the process destination within the system. The NET describes both the area and system ID of a system in the IS-IS network but excludes the NSEL, which is set to 0x00. If the NSEL identifies the process within the system, the ISO address is called the “NSAP address.”

network protocol data unit (NPDU) See *protocol data unit (PDU)*.

network selector (NSEL) In IS-IS, this is sometimes referred to as the SEL field. This field describes the service at the network layer to which the packet is to be sent. NSEL is similar to the Protocol field in IP.

network service access point (NSAP) In IS-IS, this describes a service at the network layer to which the packet is to be directed. The NSAP is the NET address with the NSEL field set to a positive value, a value other than 0x00.

originator ID This is a BGP attribute. It is an optional nontransitive attribute that is created by the route reflector. The attribute contains the router ID of the router that originated the route in the update. The purpose of this attribute is to prevent a routing loop. If the originating router receives its own update, it ignores the update.

Overload (OL) bit In IS-IS, the OL is set on an LSP if the router cannot store the entire link-state database. When other routers receive LSPs with this bit set, they will not send the router any transit traffic for fear that its routing table is incomplete. If the router is making decisions using incomplete data, its decisions might result in suboptimal paths or even routing loops. Traffic destined for the router can still be sent to the directly connected interfaces of a router transmitting the OL bit in its LSPs.

partial sequence number packet (PSNP) In IS-IS, PSNPs are sent on point-to-point links to explicitly acknowledge each LSP the router receives. A router on a broadcast subnetwork sends a PSNP requesting the LSPs it needs to synchronize its link-state database.

passive In EIGRP, an operational route is passive. If the path is lost, the router examines the topology table to find an FS. If there is an FS, it is placed in the routing table; otherwise, the router queries its neighbors, sending the route into active mode.

poison reverse Routing updates that specifically indicate that a network or subnet is unreachable, rather than implying that a network is unreachable by not including it in updates. It is used to avoid routing loops.

prefix list The prefix list is used to control how BGP learns or advertises updates. These replace distribute lists for BGP.

prefix mask The prefix mask identifies the number of bits in the subnet mask. It is written in the /xx format after the address. It is used in supernetting and route aggregation.

Priority A Cisco tool for OSPF allowing the designated router to be manually elected or, conversely, prevented from taking part in the DR/BDR election.

private addressing Private addressing is the means by which an organization can address its network without using a registered address from the Internet. This saves considerable address space on the Internet and eases restrictions within the organization. These addresses are defined in RFC 1918. They include networks 10.0.0.0/8, 172.16.0.0/12, and 192.168.0.0/16.

protocol data unit (PDU) A unit of data passed from one layer of the OSI model to the same level of the OSI model on another node. Each layer prefixes the PDU to indicate the sending OSI layer, so the network layer sends NPDU and the data-link layer sends DLPDU.

pseudonode In IS-IS, this is the LAN identifier for a broadcast subnetwork. The pseudonode makes the broadcast medium appear as a virtual router and the routers appear as connected interfaces. The adjacency to a pseudonode is represented by the system ID of the DIS plus the circuit ID of the link. The pseudonode reduces the amount of resources needed to manage the adjacencies, because the DIS generates link-state PDUs on behalf of the pseudonode. These LSPs are sent to all the connected ISs. The routers connected to the broadcast medium no longer have to maintain the adjacencies to all other routers on the medium (thus reducing memory, CPU, and bandwidth resources).

query Sent from the EIGRP router when it loses a path to a network. If there is no alternate route (feasible successor), it will send out queries to neighbors inquiring whether they have a feasible successor. This makes the route state change to active. The queries are sent reliably.

query scoping A technique where query propagation is limited to avoid Stuck in Active (SIA).

redistribution Allowing routing information discovered through one routing protocol to be distributed in the update messages of another routing protocol. It is the process of exchanging routing updates between different routing protocols. This can be done only between protocols that support the same protocol suite at Layer 3, for example, EIGRP and OSPF for TCP/IP.

Reliable Transport Protocol (RTP) In EIGRP, this is a mechanism used to determine requirements that the packets be delivered in sequence and are guaranteed.

reply A response to the query. If a router has no information to send in a reply, it will send queries to all its neighbors.

Retransmission Timeout (RTO) In EIGRP, timer calculated in reference to the SRTT. RTO determines how long the router waits for an ACK before retransmitting the packet.

route reflector This is the BGP router that ignores the BGP split-horizon rule and is configured to forward routes from other identified iBGP clients. This removes the necessity for a fully meshed iBGP network, which preserves network resources.

route reflector client A client for which a route reflector ignores the BGP split-horizon rule. It is a router that has a TCP session with its iBGP peer. It forwards routes to the route reflector, which propagates them on to other routers. The client does not have peer connections with other clients.

route reflector cluster A cluster is the group of a route reflector and clients. There can be more than one route reflector in a cluster.

route table The routing table for EIGRP. A path is moved from the topology table to the routing table when a feasible successor is identified. See also *routing table*.

routed protocol Protocol that carries data and can be routed by a router. A router must be capable of interpreting the logical internetwork as specified by that routed protocol. Examples of routed protocols include AppleTalk, DECnet, IPX, and IP.

Routeing Domain Routeing Domain is the same as the *Administrative Domain*. It defines the boundaries of a network of interconnected routers operated and managed by the same administrative group.

routing function Process of finding a path to a destination host. Routing is very complex in large networks because of the many potential intermediate destinations that a packet might traverse before reaching its destination host.

routing loop A loop in which the routing information is fed back to the originating router as if from another router. This often happens when redistribution is configured. It can lead to confusion in the network because when the originating router loses the route, it might believe that there is an alternative path. Occurs when routers have misinformation about the network and, instead of sending traffic to the destination, pass the packets between themselves in the belief that the other router knows the path.

routing protocol Protocol that accomplishes routing through the implementation of a specific routing algorithm. Examples of routing protocols include IGRP, OSPF, and RIP.

routing table Table stored in a router or some other internetworking device that keeps track of the available networks, the best paths to those destinations, and in some cases, metrics associated with those routes.

runaway congestion A condition in which the results of network congestion cause the network to generate more traffic and compound the problem.

seed metric The metric that is given to a route when it enters the routing protocol. Most routes start with a metric of 0 because they first become known to the routing protocol to which they are directly connected. However, if they are redistributed into the routing protocol, there is no starting point from which to increment the route metric. Therefore, the default metric is configured to provide a seed metric for the redistributed routes.

SEL See network selector (NSEL)

sequence number PDU (SNP) In IS-IS, SNPs are used to acknowledge the receipt of LSPs and to synchronize link-state databases.

server A node or software program that provides services to clients.

ships in the night (SIN) Routing protocols that do not interact.

silicon switching Switching based on the Silicon Switch Engine (SSE), which allows the processing of packets independent of the Silicon Switch Processor (SSP) system processor. Silicon switching provides high-speed, dedicated packet switching.

smooth round-trip time (SRTT) In EIGRP, this is the time that the router waits after sending a packet reliably to hear an acknowledgment. This is held in the neighbor table and is used to calculate the RTO.

SPF tree A tree of the topological network. It can be drawn after the SPF algorithm has been run. The algorithm prunes the database of alternative paths and creates a loop-free shortest path to all networks. The router is at the root of the network, which is perceived from its perspective.

split horizon Routing technique in which information about routes is prevented from exiting the router interface through which that information was received. Split-horizon updates are useful in preventing routing loops.

static route A route that is explicitly configured and entered into the routing table.

stub area This is an OSPF area that does not accept external summary routes. A default route is used to direct traffic to networks unknown in the stub area. This configuration allows for small routing tables and limited SPF calculation and thus a limitation on the amount of resources required.

Stuck in Active (SIA) In EIGRP, state reached when a router has sent out queries and is waiting for ACKs from all its neighbors. The route is active until all the ACKs have been received; if they do not appear after a certain time, the router is SIA for the route.

subnet mask A 32-bit number that is associated with an IP address; each bit in the subnet mask indicates how to interpret the corresponding bit in the IP address. In binary, a subnet mask bit of 1 indicates that the corresponding bit in the IP address is a network or subnet bit; a subnet mask bit of 0 indicates that the corresponding bit in the IP address is a host bit. The subnet mask then indicates how many bits have been borrowed from the host field for the subnet field. It sometimes is referred to simply as “mask.”

subnetwork The data-link layer.

subnetwork dependent layer Interfaces with the data-link layer and hides the different kinds of data-link layers from the network layer. This sublayer transmits and receives PDUs from the subnetwork, translates DLPDUs into NPDUs, and hands them to the appropriate OSI process. The subnetwork dependent layer is also responsible for creating and maintaining adjacencies through the exchange of IS-IS Hello PDUs.

subnetwork independent layer Interfaces with the transport layer and provides it with network services. It describes how CLNS creates and maintains knowledge of the network by exchanging and processing routing information so that data can be transmitted efficiently to remote destination hosts and handed to the transport layer.

subnetwork point of attachment (SNPA) The data-link layer offers two services: the physical connection to the medium and the services offered to the physical layer and network layer. The SNPA refers to these services. The SNPA address is the physical address (for example, the MAC address on a LAN).

suboptimal path A path that is not the best path. Sometimes a less desirable path is chosen.

successor The next-hop router that passes the FC. It is chosen from the FSs as having the lowest metric to the remote network.

summarization When subnets or IANA addresses are included in a larger address, where the address is written with a prefix or mask that includes the subnets. For example, 172.16.128.0/24–172.16.240.0/24 is represented by the subnet 172.16.128.0/17.

supernet A summarization of classful networks. For example, a group of Class C addresses 200.100.16.0 through 200.100.31.0 could be summarized into the address 200.100.16.0 with a mask of 255.255.240.0 (/20).

switching function Forwarding packets from an inbound interface to an outbound interface.

synchronization rule This BGP rule states that a router cannot forward a route to an eBGP peer unless the route is in its local IP routing table. This requires that the IGP and BGP routing tables are synchronized. This is to prevent BGP from advertising routes that the autonomous system cannot direct to the destination. BGP synchronization is on by default but is commonly turned off.

topology table In EIGRP, a table that contains all the paths advertised by neighbors to all the known networks. This is a list of all the successors, feasible successors, the feasible distance, the advertised distance, and the outgoing interface. DUAL acts on the topology table to determine successors and feasible successors by which to build a routing table.

triggered update See *flash update*.

Type/Length/Value (TLV) TLV is the same as a CLV, but some literature refers to the variable-length fields as Type/Length/Value in accordance with the IP terminology.

update An EIGRP packet containing change information about the network. It is sent reliably. It is sent only when there is a change in the network to affected routers.

variable-length subnet mask (VLSM) The capability to specify a different subnet mask for the same network number on different subnets. VLSM can help optimize available address space. Some protocols do not allow the use of VLSM. See also *classless routing protocols*.

virtual LAN (VLAN) A logical (rather than physical) grouping of devices into a common broadcast domain. A VLAN is defined on a LAN switch and is applied to the appropriate switch ports. This means that the devices associated with the logical network do not have to be geographically local to one another.

Index

A

ABR (area border router), 226, 270, 297

- configuring, 735–738
- summarizing, 718
- verifying, 738–746

access

- broadcast multiaccess networks, 202
- IP. *See* IP
- lists, 94
 - prioritization, 95–96*
 - security, 94*

Access layer

- design, 91
- hierarchical designs, 91

adding networks, 456

addresses

- aggregation, 112
- class addresses, 45
- exhaustion, 68
- hierarchies, 354–367
- Integrated IS-IS, 350–354
- Internet addresses, 45
- IP, 712–722
- IPv6, 108–117
- ISO, 352
- Layer 3, 41–50
- major addresses, 45
- multicast
 - IPv6, 112*
 - link-state routing protocols, 161*
- NAT, 105–108
- NET, 353–354
- network addresses, 41–45
- next-hop, 533
- private, 103–105
- supernet addresses, 45
- translating, 105–107

- unicast, 111

- VLSM addresses, 60

adjacencies, 317, 342, 357–358, 369

- broadcast links, 358
- NBMA links, 359
- neighbors, 162, 786
- OSPF, 187
- point-to-point links, 358
- troubleshooting, 317

adjacent OSPF neighbors, 187

administration

- BGP, 535–536
- fully meshed networks, 566

administrative distance

- configuring, 637
- routing protocols, 146–148

advanced distance vector protocol, EIGRP, 449

advantages

- of CIDR, 55–56
- of neighbors, 187
- of policy routing, 674–675
- of summarization, 67–68, 74
- of VLSM, 57, 60

advertisements, 532

- from BGP to IGP, 593–594
- from IGP to BGP, 593

aggregate-address command, 533, 593

aggregate global IPv6 unicast addresses, 111

aggregation

- addresses, 112
- routes, 64, 519, 533
- SLA, 113

algorithms

- Bellman Ford routing, 144
- Dijkstra, 199
- DUAL, 139, 448, 462

allocating VLSM addresses, 60

alternate paths, 458
alternatives to summarization, 71
applications, NAT, 105–108
applying BGP, 517
area area-id stub command, 302
area area-id stub no-summary command, 303
area border router (ABR), 226, 270, 297
 configuring, 735–738
 summarizing, 718
 verifying, 738–746
area default-cost command, 303
Area field, 351
area range command, 299
area virtual-link command, 305
areas, 342
 boundaries, 269
 design considerations, 385–386, 812
 interarea routing, 354
 routing, 355, 357
 types of, 272
ASBR (Autonomous System Boundary Router), 270
assigning IP VLSM subnets, 62–63
attributes
 BGP, 523
 tuning, 586–592
 verifying, 590
 local preference, 586
 next-hop, 528
 of OSPF, 162
 Originator-ID, 568
 weight, 586
autoconfiguration. *See also* configuration
 IPv6 multicast addresses, 113
auto-cost command, 231
auto-cost reference-bandwidth command, 230
automatic summarization, 68

Autonomous System Boundary Router (ASBR), 270
autonomous systems, 150, 513–518, 536. *See also* BGP
avoiding routing loops when redistributing, 624

B

backbones
 areas, 273
 routers, 270
backup designated routers, 188
bandwidth
 configuring over NBMA clouds, 491
 multipoint networks, 490
 NBMA clouds, 490
 reduction of (EIGRP), 449
bandwidth command, 230, 489
bandwidth-percent command, 491
Bellman Ford routing algorithm, 144
best path selection, Integrated IS-IS, 383–384, 811
BGP (Border Gateway Protocol), 509
 applying, 517
 attributes, 523, 528
 communications with autonomous systems, 513–518
 components, 513
 configuration commands, 530–535
 connections, 536–537
 convergence, 165
 default metrics, 635
 iBGP
 building networks, 562
 configuring, 566–574
 requirements, 562–566
 verifying, 575–577

- IGP redistribution, 593
- Internet, 583–585
- managing, 535–536
- overview of, 518–519
 - CIDR*, 519
 - policy-based routing*, 521–529
 - selection process*, 529–530
- paths, 586–592
- route aggregation, 533
- terminology, 515
- traffic, 578–582
- BGP-4 (Border Gateway Protocol version 4)**
 - internal BGP, 163
 - prefix lists, 579
 - verifying configurations, 588–590
- bit allocation**, 47
- Border Gateway Protocol**. *See* BGP
- boundaries**. *See* ABR
- boundaries, selection of**, 269
- broadcasts**
 - Frame Relay, 405
 - modes, 236
 - multiaccess networks, 202
 - next-hop attributes*, 528
 - OSPF network topologies*, 202
- building**
 - neighbor tables, 453
 - routing tables, 194
- burned-in addresses (MAC address)**, 113

C

- caches**, 27
- case studies, VLSM**, 57, 59
- CBWFQ (class-based weighted fair queuing)**, 96
- characteristics of Layer 3 addresses**, 42
- checking OSPF configurations**
 - multiarea networks, 309
 - single routers, 237
 - show ip ospf database*, 240
 - show ip ospf interface*, 242, 244
 - show ip ospf neighbor*, 244, 246
 - show ip protocols*, 247–248
 - show ip route*, 249
- CIDR (classless interdomain routing)**, 50
 - advantages of, 55–56
 - BGP overview, 519
 - definition of, 51
 - example of, 53, 73
 - example of use in case study, 53
 - IP addressing and the Internet, 52
- circuits**, 342. *See also* connections
- Cisco routers**, 26. *See also* routers
- Cisco-specific NBMA configurations**, 205
- Class addresses**, 45
- class-based weighted fair queuing (CBWFQ)**, 96
- classes, IP addresses**, 45
- classful IP routing protocols**, 11
- classful routing protocols**, 49–50
- classless interdomain routing**. *See* CIDR
- classless routing protocols**, 11–12, 138, 771
- CLNP (Connectionless Network Protocol)**, 342
- CLNS (Connectionless Network Service)**, 342
- clusters**, 569
- CLV (Code/Length/Value)**, 342
- commands**
 - aggregate-address, 533
 - area commands
 - area area-id stub*, 302
 - area area-id stub no-summary*, 303
 - area default-cost*, 303
 - area range*, 299
 - area virtual-link*, 305
 - auto-cost, 231
 - auto-cost reference-bandwidth, 230
 - bandwidth, 229–230, 489, 491
 - bandwidth-percent, 491
 - clns routing, 398
 - configuration, 530–535
 - optional*, 481–491, 824–825
 - required*, 479–481
 - cost, 229–231
 - debug, 250, 421–422, 496, 744
 - debug commands for EIGRP*, 496
 - debug ip ospf adjacency*, 195
 - default metric, 636
 - distance, 638, 647
 - distribute-list in, 650
 - distribute-list out, 650
 - EIGRP configuration, 479
 - fair-queue x y z interface configuration, 96
 - frame-relay map clns, 405
 - internal routers, 224

ip commands

- ip classless*, 12
- ip hello-interval eigrp*, 487
- ip hold-timer EIGRP*, 488
- ip ospf cost*, 230
- ip ospf network*, 232–233
- ip ospf network non-broadcast*, 236
- ip ospf network point-to-multipoint non-broadcast*, 235
- ip route*, 20
- ip unnumbered*, 68

match, 681

multiarea OSPF networks, 296–308

neighbor, 234–236

network, 226, 297, 480

no ip address, 237

no ip classless, 12

no-summary, 303

ODR configuration, 25

ospf commands

- ospf auto-cost*, 231
- ospf auto-cost reference-bandwidth*, 230
- ospf auto-cost reference-bandwidth router global configuration*, 230
- ospf neighbor*, 234

OSPF network command, 225

passive-interface, 481

prefix-list, 579

priority, 96, 188, 231

redistribute, 633, 635

redistribute connected subnets, 227

route-map, 686

router ospf process number, 225, 296

set, 631, 676, 682

show commands, 238, 407

- show clns interface*, 412, 414
- show clns neighbor*, 759
- show clns neighbor detail*, 761
- show clns neighbors*, 409–411
- show clns neighbors detail*, 410
- show interface*, 489
- show ip bgp*, 590, 759
- show ip eigrp neighbors*, 492, 743
- show ip eigrp topology*, 492–495, 742
- show ip eigrp topology all*, 492
- show ip eigrp traffic*, 492, 495
- show ip ospf*, 238, 240, 252

- show ip ospf border-routers*, 309
- show ip ospf database*, 238, 240, 252
- show ip ospf interface*, 238–244, 252
- show ip ospf neighbor*, 195, 244, 246
- show ip prefix-list*, 581
- show ip protocols*, 247–248, 743
- show ip route*, 17, 249, 744
- show ip route eigrp*, 491
- show isis database*, 414, 416, 761
- show isis database detail*, 416–419
- show isis spf-log*, 419, 421
- show memory free*, 269
- show processes cpu*, 269
- show running-config*, 724

summarization, 300

variance, 144, 485–486

communications with autonomous systems, 513–518

compatibility, IGRP, 450

components

BGP, 513

EIGRP, 450–451

composite metrics, EIGRP, 450

configuration

ABRs, 735–746

administrative distance, 638

autoconfiguration, 113

bandwidth over NBMA clouds, 491

BGP, 535–536

commands

BGP, 530–535

multiarea OSPF networks, 296–308

default metrics, 636

default static routes, 21

designated routers, 189

EIGRP

optional commands, 485, 490–491, 824–825

required commands, 479–481

floating static routes, 22

Integrated IS-IS, 397–398, 427–436

NBMA, 404, 407

routing level, 399–401

summarization, 401–403

metrics, 635

monitoring, 684–688

neighbor tables, 452

ODR, 23–25

- OSPF, 224–228, 723–724
 - broadcast mode*, 236
 - Frame Relay*, 236–237
 - in NBMA mode*, 235
 - in point-to-multipoint mode*, 236
 - in point-to-point mode on a Frame Relay subinterface*, 236
 - NBMA, 235
 - on a single router*, 232
 - on internal routers*, 225
 - options*, 228–233
 - point-to-multipoint mode*, 235–236
 - single routers*, 237–249
 - troubleshooting*, 249–250, 268–269
 - verifying*, 724–731
- prefix lists, 580
- redistribution, 632
- RID, 229
- RIPv2, 142
- route reflectors, 571
- router ID, 228
- routing tables, 460
- static routes, 20
- summarization, 723–731
- syntax, 633
- topology tables, 454
- virtual links, 306
- working configurations, 308
- connections**
 - BGP, 536–537
 - CIDR*, 519
 - controlling traffic*, 578–582
 - iBGP*, 562
 - Internet*, 583–585
 - overview of*, 518–519
 - policy-based routing*, 521–529
 - selection process*, 529–530
 - iBGP, 562
 - layer design, 91
 - NAT, 105–108
 - remote devices, 114
 - stub areas, 723–731
- content of neighbor tables, 452**
- controlling**
 - BGP traffic, 578
 - redistribution, 632
 - terminal access, 94
- convergence, 148, 784**
 - BGP, 165

- EIGRP, 150
- IGRP, 149
- IS-IS, 165
- OSPF, 165
- RIPv1, 149

conversion, RIP to EIGRP, 703

Core layer, 92

cost, 229, 231

counting to infinity, 140, 144

CSNP (complete sequence number packet), 342

custom queuing, 96

D

databases

- forwarding, 202
- interarea routing, 354
- topologies, 197

debug commands, 250, 421–422, 744

- debug ip ospf adjacency*, 195
- debug ip policy*, 684
- EIGRP, 496

debugging OSPF multiarea networks, 316.

See also troubleshooting

decision process (Integrated IS-IS), 383–384, 811

default metrics, 229, 231, 636

default networks, 647

default routes, 14

- IGRP, 642
- redistribution, 641

default static routes, 21

default-metric command, 635

defining

- network advertisements, 532
- peer-groups, 530

deleting

- paths, 458
- routes, 193–201

deployment of EIGRP, 735–738

design. *See also configuration*

- areas, 385–386, 812
- EIGRP, 468
- iBGP, 566–574
- IP networks, 99–105
- OSPF, 723–731

designated router (DR), 188

- backup designated routers, 188

- determining with priority command, 231
- electing, 189–190
- OSPF, 189–190
- priority command, 231
- purpose for, 188
- designation of OSPF, 702**
- destination address rotary translation, 107, 778**
- determining policy-based routes, 674, 840**
- devices, remote connections, 114**
- DHCP (Dynamic Host Configuration Protocol), 103**
- diagrams, 701. *See also* maps**
- Diffusing Update Algorithm. *See* DUAL**
- Dijkstra algorithm, 199**
- DIS (designated intermediate system), 343**
- disadvantages**
 - of fully meshed networks, 566
 - of policy routing, 675
- discontiguous networks, 69–70**
- discovery routes, OSPF, 196**
- distance command, 638**
- distance vector metrics, 139**
- distance vector protocols, 782**
 - EIGRP, 138
 - IGRP, 143–145, 784
 - routing loops, 139
- distribute lists, 95, 631**
- distribute-list in command, 650**
- distribute-list out command, 650**
- Distribution layer, 91**
- DNS (Domain Name System), 103**
- Domain Specific Part (DSP), 351**
- dotted decimal form, 47**
- down state, 196**
- DSP (Domain Specific Part), 351**
- DUAL (Diffusing Update Algorithm), 139, 448, 462**
- Dual IS-IS. *See* Integrated IS-IS**
- dynamic election of designated routers, 189**
- Dynamic Host Configuration Protocol. *See* DHCP**
- dynamic routes, 14**
- dynamic source address translation, 107**
 - optional commands, 481–491, 824–825*
 - required commands, 479–481*
- convergence, 150**
- default metrics, 636**
- defaults in bandwidth utilization, 489**
- distance vector protocols, 138**
- enterprise networks, 441–447**
 - components, 450–451*
 - design, 468*
 - features, 447–450*
 - operation, 451–466*
- Hello interval timers, 487**
- hold timers, 488**
- load balancing, 485**
- metrics, 460**
- optional commands, 481**
- RIP conversion, 703**
- simple redistribution, 643**
- summarization, 482**
- terminology, 851, 860**
- transitions**
 - deploying, 735–738*
 - planning, 731–734*
 - verifying, 738–746*
- troubleshooting, 496**
- tuning, 487**
- verifying, 491–496**
- verifying operations, 491**
 - debug commands, 496*
 - show ip eigrp neighbors, 492*
 - show ip eigrp topology, 493*
 - show ip eigrp traffic, 495*
- WAN optional commands, 489**
- election. *See* selection**
- enable password, 724**
- enabling OSPF routing protocol, 225, 296**
- Enhanced IGRP. *See* EIGRP**
- entering routes, 19, 193–201**
- enterprise networks, 441–447**
 - components, 450–451*
 - design, 468*
 - features, 447–450*
 - operation, 451–466*
- exchange process, 195–196**
- exchange state, 196**
- existing IP networks, designing IP addresses for, 101**
- exstart state, 196**

E

- EIGRP (Enhanced IGRP), 145**
 - configuring, 479–491, 638

extension headers, 114
 exterior routing protocols, 150
 external BGP, 163, 562

F

fast switching, 680–682
 feasible successors, 150, 465
 features
 of EIGRP, 447–450
 of NAT, 107–108, 119
 fields
 Area, 351
 ID, 351
 Logical Hop, 16
 Metric, 15
 Network, 14
 Outgoing Interface, 15
 routing tables, 14
 SEL, 351
 filtering redistribution, 650
 finite-state machines, DUAL, 462
 firewalls. *See* security
 first octet rule, 49
 floating static routes, 22, 148
 forcing next-hop addresses, 533
 formatting, 454. *See also* configuration
 Hello packets, 361–371
 IPv6 addresses, 110
 LSPs, 363–364
 neighbor tables, 452
 networks, 90
 hierarchical designs, 90–91
 scaling, 90
 NSPs, 365
 routing tables, 460
 forwarding
 databases, 202
 process (Integrated IS-IS), 385
 Frame Relay
 broadcast Integrated IS-IS, 405
 OSPF, 236–237
 subinterfaces, 236
 topologies, 205
 full state, 197, 792
 fully meshed iBGP networks, 566
 functions. *See also* commands
 routing, 25
 switching, 26

G

gateways, IGRP, 143
 guidelines
 criteria for IP network design, 90–103
 ISO addressing, 352

H

hand-off points, 139
 headers, IPv6, 114
 Hello interval timers, 487
 Hello packets
 Integrated IS-IS, 361–371
 OSPF, 191
 Hello protocol, 187, 202–207
 hiding network changes, 67
 hierarchical designs, 91. *See also* design
 Access layer, 91
 Core layer, 92
 creating, 90–91
 design rules for layers, 92
 Distribution layer, 91
 hierarchies, Integrated IS-IS, 354–367
 hold timers, 488
 holddown, RIPv1, 141, 151, 783
 hops, 16, 139
 hybrid multipoint networks, 490

I

IANA (Internet Assigned Numbers Authority), 46
 iBGP (internal BGP)
 configuring, 566–574
 networks, 562
 requirements, 562–566
 verifying, 575–577
 ICANN (Internet Corporation for Assigned Names and Numbers), 46
 ICMP (Internet Control Message Protocol), 97
 ID field, 351
 identifying neighbors, 530
 IDP (Initial Domain Part), 351
 IGP (Interior Gateway Protocol)
 redistributing between BGP, 593
 IS-IS. *See* IS-IS

IGRP (Internet Gateway Routing Protocol), 143, 145, 784

- compatibility, 450
- convergence, 149
- default metrics, 636
- distance vector routing protocols. *See* distance vector protocols

implementation. *See also* configuration

- IP access lists, 93
- IS-IS networks, 746–763
- multiple areas, 269

incremental updates

- EIGRP, 448
- link-state routing protocols, 161

init state, 196

Initial Domain Part (IDP), 351

Integrated IS-IS, 343

- adjacencies, 357–358, 369
 - broadcast links, 358*
 - NBMA links, 359*
 - point-to-point links, 358*
- areas
 - design considerations, 385–386, 812*
 - routing, 355, 357*
- broadcast Frame Relay, 405
- configuring, 397–398, 427–436
- decision process, 383–384, 811
- forwarding process, 385
- hierarchical address structures, 354–367
- ISO addressing, 350, 352
 - NETs, 352, 354*
 - NSAPs, 352*
- L1 routers, 354
- L2 routers, 354
- Level 1-2 routers, 355
- LSP formats, 363–364
- neighbors, 357
- NBMA, 404, 407, 816
- NSP formats, 365
- packet types, 361–362, 370–371
- receive process, 385
- route summarization, 387
- routing levels, 399–401
- show commands, 407
- summarization, 401–403
- TLVs, 365–367
- troubleshooting, 419–422
- update process, 379–382, 810
- verifying operation, 407–408

- versus OSPF, 346, 348–350

interarea routing, 354

interdomain routing, 51

interfaces

- bandwidth command, 229
- EIGRP, 479–491
- loopback, 228
- null. *See* null interfaces
- Outgoing Interface field, 15
- redistribution, 628
- routers, 400
- static routes, 640
- subinterfaces, 206

interior routing protocols, 150

internal BGP, 562

internal routers, 270. *See also* routers

- loopback interfaces, 228
- OSPF
 - configuring, 224*
 - options, 228–233*

Internet. *See also* connections

- BGP, 583–585
- IP addresses, 45, 47, 51, 52
- multiple connections, 583
- private addresses on, 103–105

Internet Assigned Numbers Authority. *See* IANA

Internet authoritative bodies, IP addresses, 46

Internet Corporation for Assigned Names and Numbers. *See* ICANN

Internet Gateway Routing Protocol. *See* IGRP

IP (Internet Protocol)

- access lists
 - implementing, 93*
 - uses of, 93*
- addresses, 45, 72
 - bit allocation, 47*
 - classes, 45*
 - Internet, 45–52*
 - optimizing space, 61–64*
 - planning, 712–722*
 - private, 103–105*
 - subnet masks, 47–50*
- networks
 - design criteria for, 90–103*
 - designing, 99–103*
 - private addresses, 104–105*

- routes, 403
- subnetting, 41–50
 - assigning VSLM to WAN, 62–63*
 - rules for, 48*

ip commands

- ip classless command, 12
- ip hello-interval eigrp command, 487
- ip hold-time EIGRP command, 488
- ip ospf cost, 230
- ip ospf network, 233
- ip ospf network non-broadcast, 236
- ip unnumbered, 68

IPv6 (Internet Protocol version 6), 108–117

IS-IS

- convergence, 165
- implementing, 746–763
- link-state routing protocols, 163

ISO addresses, 352

- NETs, 352–354
- NSAPs, 352

K

- keepalives, 518

L

- L1 routers, 344, 354, 858

- L1-2 routers, 344, 858

- L2 routers, 344, 354, 858

- LAN Hello packets, 362

Layer 3

- addresses
 - definition of networks, 41*
 - IP subnetting, 41–50*
 - need for, 37*
 - network characteristics of, 42*
- protocol independence (EIGRP), 449

layers

- Access, 91
- Core, 92
- design rules for, 92
- Distribution, 91
- function of, 91
- traffic, 92

- leaking local server traffic, 91

- Level 1-2 routers, 355

- link local IPv6 unicast address, 111

- link-state advertisements (LSAs), 161, 272

- link-state routing protocols, 160–164, 166

- networks, 160–161
- path selection, 161
- updating local network tables, 161

lists

- access. *See* IP
- distribute, 95
- prefixes
 - configuring, 580*
 - verifying, 581*

- LLQ (low-latency queuing), 96

load balancing

- EIGRP, 485
- RIPv1, 141, 152

- loading state, 197, 791

- local preference attribute, 586–588

- local server traffic, leaking, 91

- logic, route maps, 676

- logical AND, subnet mask IP addresses, 47

- logical connections, iBGP, 562

- Logical Hop field, 16

- longest matches, 100

- loopback interface, 228

- loopback IPv6 unicast addresses, 111

loop-free

- networks, 448
- routing tables, 449

- loops, 139–140, 160–164

- low-latency queuing (LLQ), 96

- LSAs (link-state advertisements), 161, 272

- LSPs (link-state packet), 344, 380

M

- maintenance.** *See also* troubleshooting

- redistribution, 650
- topology tables, 455
- topology databases, 199

- major addresses, 45

management

- BGP, 535–536
- fully meshed networks, 566

manual configuration

- designated routers, 189
- RID, 228

- manual summarization, 69

maps

- routes, 631
 - monitoring, 684–688*

- logic*, 676
 - topologies, 701
- match commands**, 678, 681
- match statements**, 631, 673, 686, 841
- maximum transmission unit (MTU)**, 144
- message types**, BGP, 562
- Metric field**, routing tables, 15
- metrics**
 - configuring, 635
 - distance vector metrics, 139
 - distance vector routing protocols, 139
 - EIGRP, 460
 - OSPF, 201
 - routing, 621–650
- mobility**, IPv6, 115
- modes**, configuring PSPF, 234
- modification**
 - default metrics, 229–231
 - routers, 400
 - routes, 193–201
 - routing protocols, 146–148
- monitoring**
 - networks, 407
 - policy-based routing, 684–688
- MTU (maximum transmission unit)**, 144
- multiarea OSPF networks**
 - configuration commands for, 296–308
 - troubleshooting, 315–317
 - verifying, 309–314
- multicast addresses**
 - EIGRP updates, 448
 - IPv6, 112
 - link-state routing protocols, 161
- multihoming**, 583
- multiple areas**, 269–273
 - implementing, 269
 - operations, 274
- multiple Internet connections**, 584
- multipoint networks**, configuring, 490

N

- NAT (Network Address Translation)**, 105–108
 - features of, 107–108, 119
 - IPv6, 108–117
- NBMA (nonbroadcast multiaccess) network**, 203
 - clouds, 490
 - Frame Relay, 405

- mode, 234–235
- OSPF, 205
 - choosing topologies*, 206
 - Cisco-specific*, 205
 - configuring*, 235
 - Frame Relay topologies*, 205
 - network topologies*, 203
 - RFC-complaint*, 205
 - subinterfaces*, 207
- neighbor command**, 236
- neighbors**, 357
 - identifying, 530
 - OSPF, 187
 - adjacent neighbors*, 187
 - designated routers*, 189–190
 - finding with exchange process*, 195–196
 - hello packets*, 191
 - searching, 195
 - tables
 - building*, 453
 - creating*, 452
- NETs (network entity titles)**, 344, 353, 859
- Network Address Translation**. *See* NAT
- network command**, 226, 297
 - EIGRP, 480
 - OSPF, 225–228
- Network field**, 14–15
- networks**, 41
 - addresses, 41–45
 - bit allocation*, 47
 - translating*, 105–107
 - advertisements, 532
 - broadcast multiaccess, 202
 - congestion, 90
 - creating, 90
 - hierarchical designs*, 90–91
 - scaling to reduce congestion*, 90
 - default, 647
 - definitions of, 42
 - discontiguous networks. *See* discontiguous networks
 - enterprise
 - components (EIGRP)*, 450–451
 - design (EIGRP)*, 468
 - EIGRP*, 441–447
 - features (EIGRP)*, 447–450
 - operation (EIGRP)*, 451–466
 - hiding changes, 67

- hierarchical designs, 91
 - Access layer*, 91
 - Core layer*, 92
 - Distribution layer*, 91
 - rules for each layer*, 92
- iBGP
 - building*, 562
 - configuring*, 566–574
 - requirements*, 562–566
 - verifying*, 575–577
- IP
 - design criteria for*, 90–103
 - subnetting*, 41–50
- link-state routing protocols, 160–161
- loop-free, 448
- monitoring, 407
- OSPF, 202
 - broadcast multiaccess*, 202
 - NBMA*, 203
 - point-to-multipoint*, 203
 - point-to-point*, 203
 - virtual links*, 182, 204
- point-to-multipoint, 203
- point-to-point, 203
- reducing, 98
- topologies
 - mapping*, 701
 - tables*, 456
- traffic
 - null interfaces*, 97
 - reducing*, 96
- types of, 202–207
- Next Logical Hop field**, 17
- next-hop addresses**, 533
- next-hop attribute**, 528
- no ip address**, 237
- no ip classless command**, 12
- nonbroadcast multiaccess**. *See* NBMA
- NPDUs (network protocol data units)**, 344, 859
- NSAP (network service access point)**, 344
- NSEL (network selector)**, 344
- null interfaces**, 96
 - controlling routing updates, 630
 - internet example, 97
 - intranet example, 98
 - reducing network traffic, 98
- numbering**, 225

O

- ODR (On Demand Routing)**, 23–25
- OL (Overload) bit**, 345
- Open Shortest Path First**. *See* OSPF
- open standards**, 184
- operations**
 - EIGRP, 451–466
 - multiple area OSPF, 274
- optimizing IP address spaces**, 61–64
- optional commands**, EIGRP, 481
- options**
 - cost, 231
 - internal routers, 228–231
 - OSPF, 228–233
- ordinary areas**, 272
- Originator-ID attribute**, 568
- OSPF (Open Shortest Path First)**, 184
 - checking configurations on a single router, 237
 - show ip ospf database*, 240
 - show ip ospf interface*, 242–244
 - show ip ospf neighbor*, 244–246
 - show ip protocols*, 247–248
 - show ip route*, 249
 - commands for stub areas, 302
 - configuring, 224, 723–724
 - default metric*, 635
 - internal routers*, 224–231
 - in NBMA mode*, 234–235
 - point-to-multipoint mode*, 235–236
 - verifying*, 724–731
 - virtual links*, 305
 - convergence, 165
 - customizing, 224–228
 - broadcast mode*, 236
 - Frame Relay*, 236–237
 - NBMA*, 235
 - options*, 228–233
 - point-to-multipoint mode*, 235–236
 - single routers*, 237–249
 - troubleshooting*, 249–250, 268–269
 - designating, 702
- EIGRP
 - deploying*, 735–738
 - planning*, 731–734
 - verifying*, 738–746
- fundamentals of, 193

- Hello protocol, 187
 - Integrated IS-IS, 346–350
 - link-state routing protocols, 162
 - multiarea networks, 296–308
 - troubleshooting*, 315–317
 - verifying*, 309–314
 - multiple area, 269–274
 - NBMA
 - choosing topologies*, 206
 - Cisco-specific*, 205
 - Frame Relay topologies*, 205
 - RFC-compliant*, 205
 - subinterfaces*, 207
 - neighbors, 187
 - adjacent neighbors*, 187
 - designated routers*, 189–190
 - exchange process*, 195–196
 - hello packets*, 191
 - routes, 196
 - routing, enabling, 225
 - show commands
 - show ip ospf border-routers*, 310
 - show ip ospf virtual-links*, 312–313
 - in single areas, 193–201
 - tables
 - building*, 193–194, 201
 - learning new routes*, 200–201
 - maintaining*, 199
 - topology databases*, 197
 - updating topology tables*, 199
 - terminology, 186
 - topologies, 202–207
 - broadcast multiaccess*, 202
 - databases*, 199
 - NBMA*, 203
 - point-to-multipoint*, 203
 - point-to-point*, 203
 - virtual links*, 182, 204
 - totally stubby areas, 303
 - ospf auto-cost, 231**
 - ospf neighbor command, 234**
 - Outgoing Interface field, 15**
- P**
- packets**
 - Integrated IS-IS, 361–371
 - LSP, 363–364
 - NAT, 105–108
 - NSPs, 365
 - parameters, subnets, 745**
 - passive interfaces**
 - command, 640
 - redistribution, 628
 - static routes, 640
 - passive-interface command, 481**
 - passwords, 724. See also security**
 - PAT (port address translation), 107, 777**
 - path selection**
 - paths**
 - alternate, 458
 - BGP, 586–592
 - deleting, 458
 - redundant, 642
 - selection
 - between routing protocols*, 146
 - link-state routing protocols*, 161
 - routing protocols*, 148
 - patterns, traffics, 92**
 - peer groups**
 - defining, 530
 - iBGP, 574
 - physical connections, iBGP, 562**
 - planning**
 - IP addresses, 712–722
 - transitions, 731–734
 - plug and play. See autoconfiguration**
 - point-to-multipoint mode**
 - configuring, 235
 - OSPF, 235–236
 - point-to-multipoint networks, 203**
 - point-to-point Hello packets, 362**
 - point-to-point networks, 203, 491**
 - poison reverse, split horizon with, 140**
 - poisoned routes, 149**
 - policy-based routing, 674–680**
 - BGP, 521–529
 - characteristics of, 673
 - configuring route maps, 677, 681
 - fast switching, 680–682
 - monitoring, 684–688
 - route maps, 671–676
 - port address translation. See PAT**
 - prefix lists, 579, 598–599**
 - BGP, 578
 - examples of, 580–581
 - verifying configurations, 582

prefix routing. *See* **CIDR**

prevention of loops, 140

prioritization in access lists, 96

priority command, 231

priority queuing, 96

private addresses, 103–105
 IP networks, 104–105
 IPv6, 108–117

process ID, 225, 297

processing
 fast-switching, 680–682
 prefix lists, 579

protocols
 definition of, 9
 EIGRP, 441–447
 components, 450–451
 design, 468
 features, 447–450
 metrics, 460
 operation, 451–466
 Hello, 202–207
 IGP. *See* **IS-IS**
 independence at Layer 3, 449
 link-state routing protocols. *See* **link-state routing protocols**
 routing
 distance vector, 138–142
 IPv6, 117
 link-state, 160–164
 protocols. See routing, protocols
 RTP, 487

pseudonodes, 345

PSNP (partial sequence number packet), 345

purpose of routing protocols, 9

R

ranges, private addresses, 103–105

rapid convergence, 449

receive process (Integrated IS-IS), 385

receiving LSPs, 380

redistribute connected subnets, 227

redistribution, 146, 618
 between BGP and IGP, 593
 between routing protocols, 618, 625
 commands, 633
 default metrics, 635

 examples of, 642–646
 monitoring, 684–688
 multiple protocols, 620
 network convergence, 627
 policy-based routing, 681
 problems with, 615, 624
 routing decisions that affect, 621–650
 troubleshooting, 650

reduction

 of bandwidth use (EIGRP), 449
 network traffic, 96–98
 size of routing tables, 67

redundant paths, redistribution without, 642

reliable transport protocol (RTP), 487

remote connections, 114

remote networks, alternate paths, 458

renumbering remote device connections, 114

requirements, BGP, 562–566

RFC-compliant, NBMA, 205

RID (router ID), 225

 loopback interface, 228
 manual configuration of, 228

RIP (Routing Information Protocol)

 default metrics, 635
 EIGRP, 703
 simple redistribution, 643

RIPv1 (Routing Information Protocol version 1), 140, 141, 149

RIPv2 (Routing Information Protocol version 2), 142, 149

route-map command, 677, 683, 686

router ID. *See* **RID**

router ospf process number, 225

routers

 ABR, 226, 270, 297
 ASBR, 270
 backbone, 270
 configuring, 746–763
 deleting, 458
 designated routers, 189–190. *See also*
 designated router
 interfaces, 400
 internal, 270, 228–233
 multiple area OSPF, 269–274
 OSPF
 broadcast mode, 236
 configuring, 224–228
 Frame Relay, 236–237

- NBMA, 235
 - options, 228–233
 - point-to-multipoint mode, 235–236
 - single router configuration, 237–249
 - troubleshooting, 249–250, 268–269
- routing/switching relationships, 26
- types of, 270
- updating, 197
- routes**
 - aggregation, 519, 533
 - caches, 27
 - default, 14
 - default static, 21
 - dynamic, 14
 - entering, 19, 193–201
 - floating static, 22
 - IP, 403
 - learning new, 200
 - maps, 631
 - BGP, 578
 - characteristics of, 672
 - logic, 676
 - monitoring, 684–688
 - statements, 685
 - uses for, 671
 - OSPF, 196
 - redistributing, 626
 - reflectors, 597
 - characteristics of, 569–570
 - iBGP networks, 567, 569
 - operations, 570
 - propagating updates, 571–577
 - static, 14, 19, 628
 - configuring, 20
 - passive interfaces, 640
 - summarization, 387
 - switching, 680
- routing**
 - domains, 345
 - EIGRP, 479–491. *See also* EIGRP
 - function, 25
 - IGRP, 143
 - interarea, 354
 - loops, 139
 - metrics, 621–650
 - ODR, 23–25
 - policy-based, 674–680
 - BGP, 521–529
 - fast switching, 680–682
 - monitoring, 684–688
 - protocols
 - administration distance, 146–148
 - definition of, 6, 9
 - distance vector, 138–142
 - exterior, 150
 - IGRP, 143, 145, 784
 - interior routing protocols, 150
 - IPv6, 117
 - link-state, 160–164
 - metrics, 16
 - OSPF, 162, 225, 179
 - path selection, 148
 - path selection between, 146
 - purpose of, 9
 - RIPv1, 140–141
 - types of, 11–12, 771
 - versus routed protocols, 10, 14
 - VLSM, 56
 - snapshot, 95
 - starting, 530
 - switching relationships in Cisco routers, 26
 - tables
 - building, 197
 - creating, 460
 - entering routes, 19
 - keeping current and correct, 16–17
 - loop-free, 449
 - Metric field, 15
 - Network field, 14–15
 - Next Logical Hop field, 17
 - OSPF, 193–201
 - reducing size of, 67, 462
 - ToS, 341
 - updating, 94, 631
 - versus switching, 25
- Routing Information Protocol version 1.**
 - See* RIPv1
- RTP (reliable transport protocol), 487 rules**
 - IP addresses, 50
 - policy-based routing, 523
 - VLSM, 57

S

scalability

- EIGRP, 467
- iBGP, 566–574
- multiple area OSPF, 269–273, 274
- networks, 90

searching neighbors, 195**security**

- access lists, 94
- IPv6, 115

SEL field, 351**selection**

- BGP, 529–530
- boundaries, 269
- designated routers, 189–190, 231
- paths, 161
- topologies, 206

sending LSP, 380**set command, 631, 682****shortest path first. *See* SPF****show commands**

- show clns interface command, 412, 414
- show clns neighbor command, 759
- show clns neighbor detail command, 761
- show clns neighbors command, 409, 411
- show interface, 489
- show ip bgp command, 590, 759
- show ip eigrp neighbors, 492, 743
- show ip eigrp topology, 492–493, 495, 742
- show ip eigrp topology all-links, 492
- show ip eigrp traffic, 492, 495
- show ip ospf, 238, 240, 252
- show ip ospf border-routers, 310
- show ip ospf database, 238, 240, 252
- show ip ospf interface, 238, 242–244, 252
- show ip ospf neighbor, 195, 244, 246
- show ip protocols, 247–248, 743
- show ip route, 17, 249, 744
- show ip route eigrp, 491
- show isis database command, 414, 416, 761
- show isis database detail command, 416
- show isis spf-log command, 419, 421
- show running-config command, 724

SIA (Stuck in Active), 483**single routers, 232, 237–249****site local IPv6 unicast addresses, 111****SLA (Site Level Aggregator), 113****snapshot routing, 95****SNP (sequence number PDU), 345****SNPA (subnetwork point of attachment), 346, 863****source-based transit provider selection, 674****SPF (shortest path first), 199**

- building routing tables, 201
- routing tables, 201

split horizon rule, 139–140, 144, 151**standard areas, 272****starting routing, 530****statements**

- match, 631
- neighbor, 235

static addressing, NAT, 107**static routes, 14, 19**

- configuring, 20
- passive interfaces, 640
- redistribution, 628, 640–641
- VLSM, 56

stub areas, 272

- connecting, 723–731
- OSPF commands, 303

Stuck in Active (SIA), 483**subinterfaces, 206**

- configuring point-to-point mode, 236
- OSPF, 207

subnets

- IP, 41–50, 718
- multiple area OSPF, 269–274
- parameters, 745
- subnetwork dependent layer, 346, 863

successor, feasible distance of the, 465**summarization, 64**

- ABRs, 718
- address exhaustion, 68
- advantages of, 67–68, 74
- alternatives to, 71
- automatic summarization, 68
- command, 300
- configuring, 723–731
- discontiguous networks, 69–70
- EIGRP, 482
- IP address space, 61–64, 72
- IP routes, 403
- manual summarization, 69

supernets, 45, 64**support, topologies, 449****switching**

- fast, 680–682
- function, 26

synchronization, BGP, 564–565

T

tables

- neighbor
 - building*, 453
 - creating*, 452
- routing, 13–25
 - building*, 197
 - creating*, 460
 - link-state routing protocols*, 160–164
 - loop-free*, 449
 - modifying routes*, 193–201
 - updating*, 462
- topology
 - adding networks*, 456
 - creating*, 454
 - deleting paths/routers*, 458
 - DUAL finite-state machines*, 462
 - maintaining*, 455

terminal access, controlling, 94

terminology

- BGP, 515
- EIGRP, 445, 851, 860

timers

- Hello interval, 487
- hold, 488

TLV (Type/Length/Value), 346, 365–367

topologies

- databases, 197, 199
- mapping, 701
- NBMA, 235
- OSPF, 202–207
- selection, 206
- show commands, 408
- support, 449
- tables, 458
 - adding networks*, 456
 - creating*, 454
 - DUAL finite-state machines*, 462
 - maintaining*, 455
 - routing tables*, 199

ToS (type of service), 341

totally stubby areas, 273, 303

trace, 651

traffic

- control, 94–95
- BGP, 578–582
- fast-switching, 680–682

- internal routers, 224
- leaking, 91
- patterns, 92

transitions

- EIGRP
 - deploying*, 735–738
 - planning*, 731–734
 - verifying*, 738–746
- IPv4 to IPv6, 116

translation

- addresses, 105, 107
- NAT, 105–108

triggered updates, 141–152, 161

troubleshooting, 317

- EIGRP, 496
 - design*, 467–468
 - scaling*, 467
- Integrated IS-IS, 407, 419–422
- multiarea OSPF networks, 315–317
- OSPF, 249–250, 268–269

tuning

- attributes, 586–592
- EIGRP, 487

two-way state, 196

type of service (ToS), 341

types

- of areas, 272
- of networks, 202–207
- of routing protocols, 11–12, 771
- of routes, 14
- of routers, 270

U

unequal-cost load balancing, 450

unicast addresses, 111

unspecified IPv6 unicast addresses, 111

update process (Integrated IS-IS), 381–382, 810

updating

- EIGRP, 448
- IGRP, 143
- Integrated IS-IS, 379
- link-state routing protocols, 161
- local network tables, 161
- OSPF, 225
- routers, 197, 200
- routing, 16–17, 94, 462, 631

V

variable-length subnet masks. *See* VLSM

variance command, 485–486

verification

BGP, 535–536

attributes, 590

iBGP, 575–577

BGP configurations of attributes, 592

EIGRP

debug commands, 496

show ip eigrp neighbors, 492

show ip eigrp topology, 493

show ip eigrp traffic, 495

extended ping, 651

Integrated IS-IS operation, 407

multiarea OSPF networks, 309–314

prefix lists, 581

redistribution, 650

transitions, 738–746

virtual links, 182, 204

virtual terminal line. *See* vty

VLSM (variable-length subnet masks), 56

addresses, 60

advantages of, 57, 60

case studies, 57, 59

example of, 56

routing protocols, 56

rules of, 57

static routes, 56

vty (virtual terminal line), 94

W

WAN (wide area network)

EIGRP optional commands, 489

IP VLSM subnets, 62–63

weight attribute, 586

wildcard masks, configuring, 227

