EXAM CRAM

# CCNP® and CCIE® Enterprise Core

## ENCOR 350-401


Cram Sheet


Flash Cards


Practice Tests

DONALD BACHA

# CCNP® and CCIE® Enterprise Core

## ENCOR 350-401

## Special Offers

## ENHANCE YOUR EXAM PREPARATION

### Save 70% on Complete Video Course

The *CCNP and CCIE Enterprise Core ENCOR 350-401 Complete Video Course, Complete Video Course*, available for both streaming and download, provides you with hours of expert-level instruction mapped directly to exam objectives. Put your knowledge to the test with full practice exams powered by the Pearson Test Prep practice test software, module quizzes, and more.

### Save 80% on Premium Edition eBook and Practice Test

The *CCNP and CCIE Enterprise Core ENCOR 350-401 Exam Cram Premium Edition eBook and Practice Test* provides three eBook files (PDF, EPUB, and MOBI/Kindle) to read on your preferred device and an enhanced edition of the Pearson Test Prep practice test software. You will also receive two additional practice exams with links for every question mapped to the PDF eBook.

---

**Pearson Test Prep online system requirements:**

**Browsers:** Browsers: Chrome version 73 and above, Safari version 12 and above, Microsoft Edge 44 and above.

**Devices:** Desktop and laptop computers, tablets running Android v8.0 and above or iPadOS v13 and above, smartphones running Android v8.0 and above or iOS v13 and above with a minimum screen size of 4.7". Internet access required.

**Pearson Test Prep offline system requirements:**

Windows 10, Windows 8.1; Microsoft .NET Framework 4.5 Client; Pentium-class 1 GHz processor (or equivalent); 512 MB RAM; 650 MB disk space plus 50 MB for each downloaded practice exam; access to the Internet to register and download exam databases

---

***See card insert in the back of the book*** >>
for your Pearson Test Prep activation code and special offers.

EXAM CRAM

# CCNP and CCIE Enterprise Core ENCOR 350-401 Exam Cram

**Donald Bacha**

Pearson

**CCNP and CCIE Enterprise Core ENCOR 350-401 Exam Cram**



ISBN-13: 978-0-13-689193-2

ISBN-10: 0-13-689193-4

Library of Congress Control Number: 2021924388

ScoutAutomatedPrintCode



**Editor-in-Chief**
Mark Taub

**Director, ITP Product Management**
Brett Bartow

**Executive Acquisitions Editor**
James Manly

**Development Editor**
Ellie Bru

**Managing Editor**
Sandra Schroeder

**Project Editor**
Mandie Frank

**Copy Editor**
Kitty Wilson

**Indexer**
Erika Millen

**Proofreader**
Gill Editorial Services

**Technical Editor**
Raymond Lacoste

**Publishing Coordinator**
Cindy Teeters

**Designer**
Chuti Prasertsith

**Compositor**
codeMantra

## Pearson's Commitment to Diversity, Equity, and Inclusion

Pearson is dedicated to creating bias-free content that reflects the diversity of all learners. We embrace the many dimensions of diversity, including but not limited to race, ethnicity, gender, socioeconomic status, ability, age, sexual orientation, and religious or political beliefs.

Education is a powerful force for equity and change in our world. It has the potential to deliver opportunities that improve lives and enable economic mobility. As we work with authors to create content for every product and service, we acknowledge our responsibility to demonstrate inclusivity and incorporate diverse scholarship so that everyone can achieve their potential through learning. As the world's leading learning company, we have a duty to help drive change and live up to our purpose to help more people create a better life for themselves and to create a better world.

Our ambition is to purposefully contribute to a world where

- Everyone has an equitable and lifelong opportunity to succeed through learning
- Our educational products and services are inclusive and represent the rich diversity of learners
- Our educational content accurately reflects the histories and experiences of the learners we serve
- Our educational content prompts deeper discussions with learners and motivates them to expand their own learning (and worldview)

While we work hard to present unbiased content, we want to hear from you about any concerns or needs with this Pearson product so that we can investigate and address them.

Please contact us with concerns about any potential bias at https://www.pearson.com/report-bias.html.

# Figure Credit

| | |
|---|---|
| Figure 5-1; Figure 5-2 | Courtesy of Cisco Systems, Inc. |
| Figure 5-3; Figure 5-4 | |
| Figure 5-5 | |
| Figure 5-6 | Screenshot of Monitor Section for a Cisco WLC © Cisco Systems, Inc |
| Figure 5-7 | Courtesy of Cisco Systems, Inc. |
| Figure 9-1; Figure 9-2; | Courtesy of Cisco Systems, Inc. |
| Figure 9-3; Figure 9-4; | |
| Figure 9-5; Figure 9-6; | |
| Figure 9-7; Figure 9-8; | |
| Figure 9-9; Figure 9-10; | |
| Figure 9-11; Figure 9-12; | |
| Figure 9-13 | |
| Figure 15-2; Figure 15-3; | Courtesy of Cisco Systems, Inc. |
| Figure 15-4 | |
| Figure 20-6 | Courtesy of Cisco Systems, Inc. |
| Figure 22-3 | Screenshot of The Cisco vManage Main Dashboard © Cisco Systems, Inc |
| Figure 23-1 | Screenshot of Cisco DNA Center © Cisco Systems, Inc |
| Figure 24-2 | Screenshot of Cisco WLC QoS profiles © Cisco Systems, Inc |
| Figure 26-2; Figure 26-3 | © 2022 VMware, Inc |
| Figure 26-5a; Figure 26-5b | Courtesy of Cisco Systems, Inc. |
| Figure 26-5c | |
| Figure 31-2; Figure 31-3; | Courtesy of Cisco Systems, Inc. |
| Figure 31-4 | |

# Contents at a Glance

# Table of Contents

# About the Author

**Donald Bacha** is a systems engineer with a health research organization. He's the technical lead responsible for the design and implementation of networking, compute, virtualization, storage, and disaster recovery systems. Over the past 18 years, Donald has supported cloud services provider, enterprise, and data center environments by contributing to complex routing and switching, data center, storage, and virtualization projects in both greenfield and brownfield deployments. His certifications include CCNP Enterprise, CCNP Data Center, and VCAP-DCV. He holds a master's of business administration. Donald can be found at www.allthingsvirtual.net and on Twitter at @donald_bacha.

# Dedication

*First, I dedicate this book to our Lord and Savior Jesus Christ (I can do all things through Christ which strengthens me.—Philippians 4:13). He has blessed me with the opportunity to learn, write, and share my knowledge. To my father and mother, thank you for always supporting and encouraging me.*

# Acknowledgments

A debt of gratitude goes out to executive acquisitions editor James Manly for giving me the opportunity to author this book and for his guidance. A special thank you to my development editor, Ellie Bru, who did well working to get this title out and for making it as strong as it can be. Many thanks go out to Mandie Frank and Kitty Wilson for ensuring that this book looks good and reads easily. I would like to thank the entire Pearson team and those who contributed in one way or another to this project.

# About the Technical Reviewer

**Raymond Lacoste** has dedicated his career to developing the skills of those interested in IT. In 2001, he began to mentor hundreds of IT professionals pursuing their Cisco certification dreams. This role led to teaching Cisco courses full time. Raymond is currently master instructor for Cisco Enterprise Routing and Switching, AWS, and ITIL at StormWind Studios. Raymond treats all technologies as an escape room, working to uncover every mystery in the protocols he works with. Along this journey, Raymond has passed more than 110 exams, and his office wall includes certificates from Microsoft, Cisco, ISC2, ITIL, AWS, and CompTIA. If you were visualizing Raymond's office, you'd probably expect the usual network equipment, certifications, and awards. Those certainly take up space, but they aren't his pride and joy. Most impressive, at least to Raymond, is his gemstone and mineral collection; once he starts talking about it, he just can't stop. Who doesn't get excited by a wondrous barite specimen in a pyrite matrix? Raymond presently resides with his wife and two children in eastern Canada, where they experience many adventures together.

# We Want to Hear from You!

As the reader of this book, *you* are our most important critic and commentator. We value your opinion and want to know what we're doing right, what we could do better, what areas you'd like to see us publish in, and any other words of wisdom you're willing to pass our way.

We welcome your comments. You can email or write to let us know what you did or didn't like about this book—as well as what we can do to make our books better.

*Please note that we cannot help you with technical problems related to the topic of this book.*

When you write, please be sure to include this book's title and author as well as your name and email address. We will carefully review your comments and share them with the author and editors who worked on the book.

Email: community@informit.com

# Reader Services

Register your copy of *CCNP and CCIE Enterprise Core ENCOR 350-401 Exam Cram* at www.pearsonitcertification.com for convenient access to downloads, updates, and corrections as they become available. To start the registration process, go to www.pearsonitcertification.com/register and log in or create an account*. Enter the product ISBN 9780136891932 and click Submit. When the process is complete, you will find any available bonus content under Registered Products.

*Be sure to check the box that you would like to hear from us to receive exclusive discounts on future editions of this product.

# Introduction

Welcome to *CCNP and CCIE Enterprise Core ENCOR 350-401 Exam Cram*. This book is a late-stage preparation tool that covers the CCNP/CCIE ENCOR 350-401 certification exam. It provides the information you need to quickly and efficiently go over all the topics covered on the CCNP/CCIE ENCOR 350-401 exam. This *Exam Cram* provides concise and exam-focused coverage of all of the CCNP/CCIE ENCOR 350-401 exam domains and objectives. It allows you to assess your preparedness and helps you to practice through questions and examples of the exam topics. The information you find in this *Exam Cram* will aid you in your success as you build knowledge, gain experience, and review for the CCNP/CCIE ENCOR 350-401 exam.

# About CCNP ENCOR 350-401 Exam Cram

This *Exam Cram* follows a predefined structure that makes the book easy to study as it provides the material in a concise manner. It also allows for the testing of knowledge as you go through each chapter, covering the various ENCOR domains and objectives. This book includes the following helpful elements:

- **Cram Sheet:** This foldout tear card that appears inside the front cover of the book presents important information that you should go over just before taking the exam. It is the most important "cram" element of the book and, as such, is presented as concisely as possible.
- **Chapter Topics:** Each chapter begins with a list of the exam objectives that are covered in the chapter as well as a list of the main topics in the chapters. The chapter's topics are then covered in a concise manner, with brief examples and figures where needed.
- **CramSavers:** Each chapter contains a short-answer quiz that allows you to assess how knowledgeable you are about the topics covered in the chapter. It helps you figure out if you should skip the entire chapter or skim the material and skip ahead to the Exam Alerts and CramQuizzes for particular sections.
- **Exam Alerts:** These notes provide exam-specific information that is important for you to know before you take the exam. Pay attention to Exam Alerts because the material they cover is likely to appear on the exam.

- **Cram Quizzes:** Each section of a chapter ends with a handful of multiple-choice questions that test your knowledge of the topics covered in that section. You will find the answers and explanations following each quiz.
- **Review Questions:** End-of-chapter review questions help you solidify what you have learned related to the topics for a particular chapter.

Chances are you have picked up this book in the early stage of your studies. The *Exam Cram* series was designed for late-stage study. So, unless you are very familiar with the technologies covered in the CCNP/CCIE ENCOR 350-401 exam and have considerable experience configuring and troubleshooting Cisco networks, it is highly recommended that you not use this book as your sole study resource. This *Exam Cram* is recommended for use after core knowledge has been built.

Both Cisco Press and Pearson IT Certification offer a number of CCNP/CCIE study materials to help you learn the core networking technologies covered on the CCNP/CCIE ENCOR 350-401 exam. The following highly recommended resources will help you gain core knowledge of the topics covered on the CCNP/CCIE ENCOR 350-401 exam:

- ***CCNP and CCIE Enterprise Core 350-401 Official Cert Guide*** **by Jason Gooley, Ramiro Garza Rios, Bradley Edgeworth, and David Hucaby (ISBN 978-1-58714-523-0):** This official cert guide provides in-depth coverage of the domains and objectives of the CCNP/CCIE ENCOR 350-401 exam.
- ***CCNP and CCIE Enterprise Core & CCNP Advanced Routing Portable Command Guide*** **by Patrick Gargano and Scott Empson (ISBN: 978-0-13-576816-7):** This book includes lots of configuration and verification examples to aid you in understanding the IOS commands you will encounter on the ENCOR and ENARSI exams.
- ***CCNP Enterprise Advanced Routing ENARSI 300-410 Official Cert Guide*** **by Raymond Lacoste and Brad Edgeworth (ISBN 978-1587145254):** I recommend that you read the routing-related chapters of this book (the first set of chapters, which covers EIGRP, OSPF, and BGP) to supplement your Layer 3 core knowledge.

The coauthor, Raymond Lacoste, is also the technical reviewer of this *Exam Cram*.

- **Cisco Modeling Labs (CML) Personal:** CML Personal (formerly Cisco VIRL) is a powerful network virtualization and orchestration platform you can use to study for Cisco certifications. CML Personal uses real Cisco IOS images and gives you the ability to simulate networks reliably. Both IOSv and IOSvL2 images are included. The majority of the topics that are coved in the CCNP/CCIE ENCOR 350-401 exam can be practiced using CML Personal. CML Personal allows up to 20 concurrent simulated nodes, and CML Personal Plus supports up to 40 concurrent simulated nodes. The majority of the examples in this *Exam Cram* were created using CML Personal. For more information on CML Personal, see https://developer.cisco.com/docs/modeling-labs. Cisco CML Personal can be purchased from the Cisco Learning Network Store at https://learningnetworkstore.cisco.com/cisco-modeling-labs-personal/cisco-cml-personal.

# About the ENCOR 350-401 Exam

The material in this *Exam Cram* closely follows the official exam domains and objectives to ensure your success on the CCNP/CCIE ENCOR 350-401 exam. To earn the CCNP Enterprise certification, there is no formal prerequisite, although Cisco recommends that you have a good understanding of the exam topics before taking the exams. In addition, Cisco recommends that CCNP candidates have three to five years of experience implementing enterprise networking solutions.

To earn the CCNP Enterprise certification, you have to pass two exams: one required exam that covers core enterprise technologies and one enterprise concentration exam of your choice, based on your technical area of focus. Passing any of these concentration exams also allows you to earn an individual Specialist certification that helps recognize your accomplishments along the way to earning your CCNP Enterprise certification. These are the requirements for earning the CCNP Enterprise certification:

- Required exam: 350-401: Implementing and Operating Cisco Enterprise Network Core Technologies (ENCOR)
- One concentration exam:

- 300-410: Implementing Cisco Enterprise Advanced Routing and Services (ENARSI)
- 300-415: Implementing Cisco SD-WAN Solutions (ENSDWI)
- 300-420: Designing Cisco Enterprise Networks (ENSLD)
- 300-425: Designing Cisco Enterprise Wireless Networks (ENWLSD)
- 300-430: Implementing Cisco Enterprise Wireless Networks (ENWLSI)
- 300-435: Implementing Automation for Cisco Enterprise Solutions (ENAUI)

This book focuses on the required 350-401 (ENCOR) exam. It is a 120-minute exam that tests your knowledge of enterprise infrastructure, including dual-stack architecture, virtualization, infrastructure, network assurance, security, and automation. The CCNP/CCIE ENCOR 350-401 exam is also the qualifying exam for the CCIE Enterprise Infrastructure and CCIE Enterprise Wireless certifications. Once you pass the CCNP/CCIE ENCOR 350-401 exam, you are automatically qualified to schedule and take the CCIE lab exam in those tracks.

# Cisco ENCOR 350-401 Exam Topics

Table I-1 lists general exam topics (that is, objectives) and specific topics under each general topic (that is, subobjectives) for the CCNP/CCIE ENCOR 350-401 exam. This table also lists the chapter in which each exam topic is covered.

This *Exam Cram* covers every domain and objective of the CCNP/CCIE ENCOR 350-401 exam. It follows the official exam objectives closely to ensure your success on the CCNP/CCIE ENCOR 350-401 exam. As such, all of the contents, including CramSaver, Cram Quizzes, and Review Questions, map to specific objectives of the CCNP/CCIE ENCOR 350-401 exam. The latest CCNP/CCIE ENCOR 350-401 exam objectives can be found on the Cisco Learning Network at https://learningnetwork.cisco.com/s/encor-exam-topics.

TABLE I-1 **ENCOR 350-401 Exam Topics**

| Chapter | ENCOR Exam Objectives |
|---|---|
| | *1.0 Architecture* |
| | 1.1 Explain the different design principles used in an enterprise network |
| 19: Enterprise Network Design Principles | 1.1.a Enterprise network design such as Tier 2, Tier 3, and Fabric Capacity planning |
| 19: Enterprise Network Design Principles | 1.1.b High availability techniques such as redundancy, FHRP, and SSO |
| | 1.2 Analyze design principles of a WLAN deployment |
| 20: Wireless LAN Deployments | 1.2.1 Wireless deployment models (centralized, distributed, controller-less, controller based, cloud, remote branch) |
| 20: Wireless LAN Deployments | 1.2.b Location services in a WLAN design |
| 21: On-Premises vs. Cloud Infrastructure | 1.3 Differentiate between on-premises and cloud infrastructure deployments |
| | 1.4 Explain the working principles of the Cisco SD-WAN solution |
| 22: SD-WAN | 1.4.a SD-WAN control and data planes elements |
| 22: SD-WAN | 1.4.b Traditional WAN and SD-WAN solutions |
| | 1.5 Explain the working principles of the Cisco SD-Access solution |
| 23: SD-Access | 1.5.a SD-Access control and data planes elements |
| 23: SD-Access | 1.5.b Traditional campus interoperating with SD-Access |
| | 1.6 Describe concepts of wired and wireless QoS |
| 24: QoS | 1.6.a QoS components |
| 24: QoS | 1.6.b QoS policy |
| | 1.7 Differentiate hardware and software switching mechanisms |
| 25: Switching | 1.7.a Process and CEF |
| 25: Switching | 1.7.b MAC address table and TCAM |
| 25: Switching | 1.7.c FIB vs. RIB |
| | *2.0 Virtualization* |
| | 2.1 Describe device virtualization technologies |
| 26: Basic Virtualization | 2.1.a Hypervisor type 1 and 2 |
| 26: Basic Virtualization | 2.1.b Virtual machine |
| 26: Basic Virtualization | 2.1.c Virtual switching |

| Chapter | ENCOR Exam Objectives |
|---|---|
| | 2.2 Configure and verify data path virtualization technologies |
| 27: VRF Instances, GRE, and IPsec | 2.2.a VRF |
| 27: VRF Instances, GRE, and IPsec | 2.2.b GRE and IPsec tunneling |
| | 2.3 Describe network virtualization concepts |
| 28: Extending the Network Virtually | 2.3.a LISP |
| 28: Extending the Network Virtually | 2.3.b VXLAN |
| | *3.0 Infrastructure* |
| | 3.1 Layer 2 |
| 1: Understanding Layer 2 | 3.1.a Troubleshoot static and dynamic 802.1q trunking protocols |
| 1: Understanding Layer 2 | 3.1.b Troubleshoot static and dynamic EtherChannels |
| 1: Understanding Layer 2 | 3.1.c Configure and verify common Spanning Tree Protocols (RSTP and MST) |
| | 3.2 Layer 3 |
| 2: Understanding Layer 3: IGPs | 3.2.a Compare routing concepts of EIGRP and OSPF (advanced distance vector vs. link state, load balancing, path selection, path operations, metrics) |
| 2: Understanding Layer 3: IGPs | 3.2.b Configure and verify simple OSPF environments, including multiple normal areas, summarization, and filtering (neighbor adjacency, point-to-point and broadcast network types, and passive interface) |
| 3: Understanding Layer 3: BGP | 3.2.c Configure and verify eBGP between directly connected neighbors (best path selection algorithm and neighbor relationships) |
| | 3.3 Wireless |
| 5: Enterprise Wireless | 3.3.a Describe Layer 1 concepts, such as RF power, RSSI, SNR, interference noise, band and channels, wireless client devices capabilities |
| 5: Enterprise Wireless | 3.3.b Describe AP modes and antenna types |
| 5: Enterprise Wireless | 3.3.c Describe access point discovery and join process (discovery algorithms, WLC selection process) |
| 5: Enterprise Wireless | 3.3.d Describe the main principles and use cases for Layer 2 and Layer 3 roaming |

| Chapter | ENCOR Exam Objectives |
|---|---|
| 5: Enterprise Wireless | 3.3.e Troubleshoot WLAN configuration and wireless client connectivity issues |
| | 3.4 IP Services |
| 4: IP Services | 3.4.a Describe Network Time Protocol (NTP) |
| 4: IP Services | 3.4.b Configure and verify NAT/PAT |
| 4: IP Services | 2.4.c Configure first hop redundancy protocols, such as HSRP and VRRP |
| 4: IP Services | 3.4.d Describe multicast protocols, such as PIM and IGMP v2/v3 |
| | *4.0 Network Assurance* |
| 29: Troubleshooting | 4.1 Diagnose network problems using tools such as debugs, conditional debugs, trace route, ping, SNMP, and syslog |
| 30: Monitoring | 4.2 Configure and verify device monitoring using syslog for remote logging |
| 30: Monitoring | 4.3 Configure and verify NetFlow and Flexible NetFlow |
| 30: Monitoring | 4.4 Configure and verify SPAN/RSPAN/ERSPAN |
| 31: IP SLA and DNA Center | 4.5 Configure and verify IPSLA |
| 31: IP SLA and DNA Center | 4.6 Describe Cisco DNA Center workflows to apply network configuration, monitoring, and management |
| 32: NETCONF and RESTCONF | 4.7 Configure and verify NETCONF and RESTCONF |
| | *5.0 Security* |
| | 5.1 Configure and verify device access control |
| 6: Device Access Control | 5.1.a Lines and password protection |
| 6: Device Access Control | 5.1.b Authentication and authorization using AAA |
| | 5.2 Configure and verify infrastructure security features |
| 7: Infrastructure Security | 5.2.a ACLs |
| 7: Infrastructure Security | 5.2.b CoPP |
| 8: Securing REST APIs | 5.3 Describe REST API security |
| | 5.4 Configure and verify wireless security features |
| 9: Wireless Security | 5.4.a EAP |
| 9: Wireless Security | 5.4.b WebAuth |
| 9: Wireless Security | 5.4.c PSK |
| | 5.5 Describe the components of network security design |

| Chapter | ENCOR Exam Objectives |
|---|---|
| 10: Network Security Design | 5.5.a Threat defense |
| 10: Network Security Design | 5.5.b Endpoint security |
| 10: Network Security Design | 5.5.c Next-generation firewall |
| 10: Network Security Design | 5.5.d TrustSec, MACsec |
| 11: Network Access Control | 5.5.e Network access control with 802.1X, MAB, and WebAuth |
| | *6.0 Automation* |
| 12: Anatomy of Python | 6.1 Interpret basic Python components and scripts |
| 13: Building JSON Files | 6.2 Construct valid JSON encoded file |
| 14: YANG Data Modeling | 6.3 Describe the high-level principles and benefits of a data modeling language, such as YANG |
| 15: DNA Center and vManage APIs | 6.4 Describe APIs for Cisco DNA Center and vManage |
| 16: Interpreting REST API Codes | 6.5 Interpret REST API response codes and results in payload using Cisco DNA Center and RESTCONF |
| 17: EEM Applets | 6.6 Construct EEM applet to automate configuration, troubleshoot, or data collection |
| 18: Configuration Management and Orchestration | 6.7 Compare agent vs. agentless orchestration tools, such as Chef, Puppet, Ansible, and SaltStack |

# Booking and Taking the ENCOR 350-401 Exam

Because this *Exam Cram* is a late-stage study material, by the time you are reading it, you have probably already registered to take the CCNP/CCIE ENCOR 350-401 exam. If not, my recommendation is that you go ahead and register and use that registration as motivation to prepare for the exam. If you find yourself not feeling fully prepared, or if some other circumstance comes up before the exam, you can cancel your registration. Pearson Vue allows you to cancel your registration up until 24 hours before you are scheduled to take the exam without a penalty.

At this writing, Pearson Vue allows you to take the exam at one of its testing sites or from home using the OnVUE online proctoring system, where a live proctor monitors you through the webcam of your computer. If you are using the online

proctoring system, you should run the system test and exam simulation before registering. You can register online at https://home.pearsonvue.com/cisco, over the phone, or as a Pearson Vue walk-in, where available. In the United States and Canada, you can schedule your exam up to six weeks in advance, and you must wait five calendar days from the end of your first attempt before retaking the same exam. Hopefully, with the help of this *Exam Cram* and the other recommended resources, you will not have to worry about that!

To register for the exam, you need the following information:

- Legal name (from a government-issued ID)
- Cisco certification ID (for example, CSCO00000001) or test ID number
- Valid email address
- Method of payment

At this writing, the cost of the CCNP/CCIE ENCOR 350-401 exam is US$400.

# What to Expect from the Exam

If you haven't taken a certification test, the process can be a little unnerving. Even if you've taken numerous tests, it is not much better. Mastering the inner mental game often can be as much of a battle as knowing the material. Knowing what to expect before heading in can make the process a little more comfortable.

Certification tests are administered on a computer system at a VUE authorized testing center. The format of the exams is straightforward: Each question has several possible answers to choose from. The questions in this book provide a good example of the types of questions you can expect on the exam. If you are comfortable with them, the test should hold few surprises.

As you take the CCNP/CCIE ENCOR 350-401 exam, be sure to review each answer before moving on to the next question. After you answer a question, you cannot go back at a later time to make changes.

You can expect to see several types of questions on the ENCOR exam:

- **Multiple-choice, single answer:** This type of question requires you to choose only one answer for a question. Once you select the radio button for your answer, click Next to move on to another question.
- **Multiple-choice, multiple answers:** This type of question shows you how many answers you need to select. To select the answers, you click the

small squares next to the answers of your choice to insert checkmarks. Once you choose the correct number of questions, you can click Next to move on to the next question.

- **Drag and drop:** This type of question requires you to select an option on the left and drag and drop it to its appropriate drop zone on the right. Sometimes only some of the options on the left are used.
- **Fill-in-the-blank:** This type of question requires you to insert your answer in a text box. Sometimes you may have to fill in multiple text boxes.
- **Testlet:** This type of question is scenario based. It involves reading a scenario and then answering the question(s) related to the scenario. Testlet questions are typically some variation of multiple-choice questions.

Cisco has published two exam tutorial videos that provide a walk-through demonstration on the various exam question types and how they function. You can find these short videos at https://learningnetwork.cisco.com/s/certification-exam-tutorials.

# A Few Exam-Day Details

It is recommended that you arrive at the examination room at least 15 minutes early, although a few minutes earlier certainly would not hurt. This will give you time to prepare and will give the test administrator time to answer any questions you might have before the test begins. Many people suggest that you review the most critical information about the test you're taking just before the test. (Exam Cram books provide a reference—the Cram Sheet, located inside the front of this book—that lists the essential information from the book in distilled form.) Arriving a few minutes early will give you some time to compose yourself and mentally review this critical information.

You will be asked to provide two forms of ID, one of which must be a photo ID. Both of the forms of ID you choose should have signatures. You also might need to sign in when you arrive and sign out when you leave.

Be warned: The rules are clear about what you can and cannot take into the examination room. Books, laptops, note sheets, and so on are not allowed in the examination room. The test administrator will hold these items, to be returned after you complete the exam. You might receive either a wipe board or a pen and a single piece of paper for making notes during the exam. The test administrator will ensure that no paper is removed from the examination room.

# After the Test

Whether you want it or not, as soon as you finish your test, your score displays on the computer screen. In addition to the results appearing on the computer screen, a hard copy of the report prints for you. Like the onscreen report, the hard copy displays the results of your exam and provides a summary of how you did on each section and on each technology. If you were unsuccessful, this summary can help you determine the areas you need to brush up on. After you have taken the CCNP/CCIE ENCOR 350-401 exam, please note the following:

- Every written proctored exam passed equals a Specialist certification.
- Within 24 hours of passing your certifying exam, you will receive an email advising you on the next steps. You must complete the steps to trigger the fulfillment process.
- The Cisco Certification Tracking System records exam and certification status. Be sure to keep your contact information up to date if you want to receive notifications.
- After you're certified, you will be authorized to use the Cisco Certification logo that identifies your status, provided that you read and acknowledge the Cisco Certifications Logo Agreement. You can download logos through the Certifications Tracking System.
- Visit the Certification and Fulfillment Benefits page to learn more about the certification fulfillment process and the benefits you'll receive.

# Last-Minute Exam Tips

Studying for a certification exam is no different than studying for any other exam, but a few hints and tips can give you the edge on exam day:

- **Read all the material:** Read each question carefully and entirely before answering.
- **Watch for the Exam Alerts:** The CCNP/CCIE ENCOR 350-401 exam objectives include a wide range of technologies. Exam Alerts found throughout each chapter of this book are designed to highlight exam-related hot spots. Skim the book for Exam Alerts when preparing for the exam.
- **Use the questions to assess your knowledge:** Don't just read the chapter content; use the CramSaver questions to find out what you know and what you don't. If you struggle to answer any of these questions, read the entire chapter, including Exam Alerts, and complete the Cram Quiz

at the end of each section and the Review Questions at the end of the chapter.

- **Review the exam objectives:** Develop your own questions and examples for each topic listed. If you can develop and answer several questions for each topic, you should not find it difficult to pass the exam.

Good luck with your CCNP/CCIE ENCOR 350-401 exam studies, and thank you for selecting the *CCNP and CCIE Enterprise Core ENCOR 350-401 Exam Cram*.

# Companion Website

Register this book to get access to the Pearson Test Prep practice test software and other study materials plus additional bonus content. Check this site regularly for new and updated postings written by the author that provide further insight into the more troublesome topics on the exams. Be sure to check the box that you would like to hear from us to receive updates and exclusive discounts on future editions of this product or related products.

To access this companion website, follow these steps:

1. Go to **www.pearsonITcertification.com/register** and log in or create a new account.
2. Enter the ISBN **9780136891932**.
3. Answer the challenge question as proof of purchase.
4. Click the **Access Bonus Content** link in the Registered Products section of your account page, to be taken to the page where your downloadable content is available.

Please note that many of our companion content files, especially image and video files, can be very large.

If you are unable to locate the files for this title by following these steps, please visit www.pearsonITcertification.com/contact and select the Site Problems/Comments option. Our customer service representatives will assist you.

# Pearson Test Prep Practice Test Software

This book comes complete with the Pearson Test Prep practice test software, containing two full exams. These practice tests are available to you either online or in an offline Windows application. To access the practice exams that

were developed with this book, please see the instructions in the card inserted in the sleeve in the back of the book. This card includes a unique access code that enables you to activate your exams in the Pearson Test Prep practice test software.

> **Note**
>
> The cardboard sleeve in the back of this book includes a piece of paper. The paper lists the activation code for the practice exams associated with this book. Do not lose the activation code. On the opposite side of the paper from the activation code is a unique, one-time-use coupon code for the purchase of the Premium Edition eBook and Practice Test.

# Accessing the Pearson Test Prep Software Online

The online version of this software can be used on any device with a browser and connectivity to the Internet, including desktop machines, tablets, and smartphones. To start using your practice exams online, simply follow these steps:

1. Go to **http://www.PearsonTestPrep.com**.
2. Select **Pearson IT Certification** as your product group.
3. Enter your email and password for your account. If you don't have an account on PearsonITCertification.com or CiscoPress.com, you need to establish one by going to **PearsonITCertification.com/join**.
4. In the **My Products** tab, click the **Activate New Product** button.
5. Enter the access code printed on the insert card in the back of your book to activate your product. The product is then listed in your My Products page.
6. Click the Exams button to launch the exam settings screen and start the exam.

# Accessing the Pearson Test Prep Software Offline

If you wish to study offline, you can download and install the Windows version of the Pearson Test Prep software. You can find a download link for this software on the book's companion website, or you can just enter this link in your browser:

http://www.pearsonitcertification.com/content/downloads/pcpt/engine.zip

To access the book's companion website and the software, simply follow these steps:

1. Register your book by going to PearsonITCertification.com/register and entering the ISBN **9780136891932**.
2. Respond to the challenge questions.
3. Go to your account page and select the **Registered Products** tab.
4. Click on the **Access Bonus Content** link under the product listing.
5. Click the **Install Pearson Test Prep Desktop Version** link in the Practice Exams section of the page to download the software.
6. When the software finishes downloading, unzip all the files onto your computer.
7. Double-click the application file to start the installation, and follow the onscreen instructions to complete the registration.
8. When the installation is complete, launch the application and click the **Activate Exam** button on the My Products tab.
9. Click the **Activate a Product** button in the Activate Product Wizard.
10. Enter the unique access code from the card in the sleeve in the back of your book and click the **Activate** button.
11. Click **Next** and then click the **Finish** button to download the exam data to your application.
12. You can now start using the practice exams by selecting the product and clicking the **Open Exam** button to open the exam.

Note that the offline and online versions sync together, so saved exams and grade results recorded on one version will be available to you in the other version as well.

# Customizing Your Exams

When you are in the exam settings screen, you can choose to take exams in one of three modes:

- Study mode
- Practice Exam mode
- Flash Card mode

Study mode allows you to fully customize an exam and review answers as you are taking the exam. This is typically the mode you use first to assess your knowledge and identify information gaps. Practice Exam mode locks certain customization options in order to present a realistic exam experience. Use this mode when you are preparing to test your exam readiness. Flash Card mode strips out the answers and presents you with only the question stem. This mode is great for late-stage preparation, when you really want to challenge yourself to provide answers without the benefit of seeing multiple-choice options. This mode does not provide the detailed score reports that the other two modes provide, so it is not the best mode for helping you identify knowledge gaps.

In addition to these three modes, you will be able to select the source of your questions. You can choose to take exams that cover all of the chapters, or you can narrow your selection to just a single chapter or the chapters that make up specific parts in the book. All chapters are selected by default. If you want to narrow your focus to individual chapters, simply deselect all the chapters and then select only those on which you wish to focus in the Objectives area.

You can also select the exam banks on which to focus. Each exam bank comes complete with a full exam of questions that cover topics in every chapter. The two exams printed in the book are available to you, as are two additional exams of unique questions. You can have the test engine serve up exams from all four banks or just from one individual bank by selecting the desired banks in the exam bank area.

You can make several other customizations to your exam from the exam settings screen, such as the time of the exam, the number of questions, whether to randomize questions and answers, whether to show the number of correct answers for multiple answer questions, or whether to serve up only specific types of questions. You can also create custom test banks by selecting only questions that you have marked or questions on which you have added notes.

## Updating Your Exams

If you are using the online version of the Pearson Test Prep software, you should always have access to the latest version of the software as well as the exam data. If you are using the Windows desktop version, every time you launch the software, it will check to see if there are any updates to your exam data and automatically download any changes made since the last time you used the software. This requires you to be connected to the Internet at the time you launch the software.

Sometimes, due to a number of factors, the exam data might not fully download when you activate your exam. If you find that figures or exhibits are missing, you might need to manually update your exams.

To update a particular exam you have already activated and downloaded, simply select the **Tools** tab and click the **Update Products** button. Again, this is only an issue with the desktop Windows application.

If you want to check for updates to the Windows desktop version of the Pearson Test Prep exam engine software, simply select the **Tools** tab and click the **Update Application** button. Doing so allows you to ensure that you are running the latest version of the software engine.

## Assessing Exam Readiness

Exam candidates never really know whether they are adequately prepared for the exam until they have completed about 30% of the questions. At that point, if you are not prepared, it is too late. The best way to determine your readiness is to work through the CramSaver quiz at the beginning of each chapter and review the topics presented in each chapter. It is best to work your way through the entire book unless you can complete each subject without having to do any research or look up any answers.

## Premium Edition eBook and Practice Tests

This book includes an exclusive offer for 70% off the Premium Edition eBook and Practice Tests edition of this title. Please see the coupon code included with the cardboard sleeve for information on how to purchase the Premium Edition.

CHAPTER 1

# Understanding Layer 2

**This chapter covers the following official ENCOR 350-401 exam objectives:**

- 3.1 Layer 2
- 3.1.a Troubleshoot static and dynamic 802.1q trunking protocols
- 3.1.b Troubleshoot static and dynamic EtherChannels
- 3.1.c Configure and verify common Spanning Tree Protocols (RSTP and MST)

This chapter is divided into three sections. It covers a host of Layer 2 technologies and is one of the longest chapters in this book.

The first section covers the features that are necessary for switch-to-switch connectivity. It starts with an overview of the configuration and troubleshooting of VLANs and 802.1Q trunking technologies. It looks at how VLAN Trunking Protocol (VTP) and Dynamic Trunking Protocol (DTP) assist in the provision of VLANs and the carrying of VLANs across switches, respectively. Next, this section looks at inter-VLAN routing using a router and a Layer 3 switch.

The second section of this chapter covers the Layer 2 loop avoidance mechanism Spanning Tree Protocol. It looks at the configuration and verification of Cisco Per VLAN Spanning Tree (PVST/PVST+), Rapid Spanning Tree Protocol (RSTP), Multiple Spanning Tree Protocol (MST), and the various Spanning Tree Protocol tuning and protection mechanisms.

The final section of this chapter covers configuration and troubleshooting of Layer 2 and Layer 3 EtherChannels.

**This chapter covers the following technology topics:**

- VLANs Overview
  - 802.1Q Trunking
  - VLAN Trunking Protocol (VTP)
  - Dynamic Trunking Protocol (DTP)
  - Inter-VLAN Routing
- Spanning Tree Protocol Overview
  - Rapid Spanning Tree Protocol (RSTP)
  - Spanning Tree Protocol (STP) Tuning and Protection Mechanisms
  - Multiple Spanning Tree Protocol (MST)
- EtherChannels

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. What is the range of extended VLANs?
2. What VTP version is used by default on Cisco Catalyst switches?
3. Which spanning-tree protocol allows for interoperability between Cisco and other vendors switches?
4. Which spanning-tree protection mechanism works by preventing BPDUs from being sent out?
5. What command is used to actively negotiate the forming of an LACP EtherChannel?
6. Which EtherChannel aggregation protocol is Cisco proprietary and thus only works between Cisco devices?

### Answers

1. 1006 through 4094
2. VTP Version 1
3. Multiple Spanning Tree (MST)
4. BPDU Filter
5. **channel-group** *number* **mode** *active*
6. PAgP

# VLANs Overview

This section of the chapter looks at the configuration and troubleshooting of VLANs and 802.1Q trunking technologies. It looks at how VLAN Trunking Protocol (VTP) and Dynamic Trunking Protocol (DTP) assist in the provision of VLANs and carrying of VLANs across switches, respectively. This section also looks at inter-VLAN routing using a router and a Layer 3 switch.

A virtual local area network (VLAN) is a group of end stations in a switched network that is logically segmented by function or application, without regard to the physical locations of the users. VLANs have the same attributes as physical LANs, but you can group end stations even if they are not physically located on the same LAN segment.

Any switch port can belong to a VLAN, and unicast, broadcast, and multicast packets are forwarded and flooded only to end stations in that VLAN. Each VLAN is considered a logical network, and packets destined for stations that do not belong to the VLAN must be forwarded through a router or a Layer 3 switch. VLANs are usually associated with IP subnetworks. For example, all the end stations in a particular IP subnet belong to the same VLAN. To allow VLAN-to-VLAN communication, the traffic needs to go to a router (router-on-a-stick) or must be routed between switch virtual interfaces (SVIs) of a Layer 3 switch.

By default, a newly created VLAN is operational and in the no shutdown condition. In addition, you can configure VLANs to be in the active state, where they pass traffic, or the suspended state, in which they do not pass traffic. By default, VLANs are in the active state and pass traffic.

VLANs are numbered from 1 to 4094. VLANs in the normal range are identified with a number from 1 to 1001. VLANs 1002 through 1005 are reserved; these VLANs are automatically created and cannot be removed. VLANs between 1006 and 4094 are part of the extended VLAN range. All ports that you configure as switch ports belong to the default VLAN. The default VLAN (VLAN 1) uses only default values, and you cannot create, delete, or suspend activity in the default VLAN.

You create a VLAN by assigning a number to it; you can delete VLANs and move them from the active operational state to the suspended operational state. If you attempt to create a VLAN with an existing VLAN ID, the device goes into the VLAN sub-mode but does not create the same VLAN again.

# VLAN Assignment

A newly created VLAN remains unused until Layer 2 ports are assigned to that specific VLAN. All the ports are assigned to VLAN 1 by default. Depending on the range of the VLAN, you can configure the following parameters for VLANs (except the default VLAN):

- VLAN name
- VLAN state
- Shutdown or no shutdown

When you delete a specified VLAN, the ports associated with that VLAN become inactive, and no traffic flows. When you delete a VLAN from a trunk port (as discussed next), only that VLAN is shut down, and traffic continues to flow for all the other VLANs through the trunk port.

However, when you delete a VLAN, the system retains all the VLAN-to-port mapping for that VLAN, and when you re-enable or re-create that specified VLAN, the system automatically reinstates all the original ports to that VLAN. However, the static MAC addresses and aging time for that VLAN are not restored when the VLAN is re-enabled.

Next, let us review the assignment and verification of access ports to a particular VLAN. Assigning a switch port to an access VLAN allows for the forwarding of frames from a port in a particular VLAN to another port in the same VLAN.

You create a VLAN by using the command **vlan** *vlan-id*, which gets you into the VLAN configuration mode.

Example 1.1 shows how to create two VLANs (10 and 20). You can optionally enter a name for a VLAN by using the command **vlan** *name*. In this case, VLAN 10 and VLAN 20 are named Finance and Sales, respectively. You can verify the creation of VLANs by using the command **show vlan brief**.

EXAMPLE 1.1 **Creating and Verifying VLANs**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# vlan 10
SW1(config-vlan)# name Finance
SW1(config-vlan)# vlan 20
SW1(config-vlan)# name Sales
SW1(config-vlan)# end
```

```
SW1#
SW1# show vlan brief

VLAN Name                             Status    Ports
---- -------------------------------- --------- -------------------------
1    default                          active    Gi0/0, Gi0/1, Gi0/2, Gi0/3
10   Finance                          active
20   Sales                            active
1002 fddi-default                     act/unsup
1003 token-ring-default               act/unsup
1004 fddinet-default                  act/unsup
1005 trnet-default                    act/unsup
SW1#
```

After you create a VLAN, you can associate a switch port with that VLAN. At that point, you can also specify that the switch port mode is an access port. An access port is associated with one VLAN. When you specify that a port is configured in access mode, you are telling the switch that it should carry one VLAN. By default, this is VLAN 1. A VLAN can also be created at the point where a switch port is being assigned to a VLAN. If a VLAN is not yet defined on a Cisco Catalyst switch, once you assign a switch port to a VLAN, that VLAN will automatically be created, and the switch port will be placed inside it.

Example 1.2 shows the use of the command **switchport mode access** to assign a switch port as an access port. The command **switchport access vlan** *vlan-id* places the switch port into a specific VLAN (VLANs 10 and 20, respectively, in this case).

EXAMPLE 1.2 **Assigning Ports to a VLAN**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# interface GigabitEthernet 0/1
SW1(config-if)# switchport mode access
SW1(config-if)# switchport access vlan 10
SW1(config-if)# interface GigabitEthernet 0/2
SW1(config-if)# switchport mode access
SW1(config-if)# switchport access vlan 20
SW1(config-if)# end
SW1#
```

The command **show vlan brief** shows a single line of output for each VLAN, with the name of the VLAN, its status, and the ports that are assigned to the VLANs.

Example 1.3 shows the output of the **show vlan brief** command. It shows the default VLAN, VLAN 1, as well as VLANs 10 and 20, which are defined in Example 1.2. It also shows the status of those VLANs as active and the access ports that are assigned to those VLANs.

EXAMPLE 1.3 **Verifying a VLAN by Using the show vlan brief Command**

```
SW1# show vlan brief

VLAN Name                              Status    Ports
---- --------------------------------- --------- -------------------
1    default                           active    Gi0/0, Gi0/3
10   Finance                           active    Gi0/1
20   Sales                             active    Gi0/2
1002 fddi-default                      act/unsup
1003 token-ring-default                act/unsup
1004 fddinet-default                   act/unsup
1005 trnet-default                     act/unsup
SW1#
```

You can verify the access port configuration of a switch port by using the command **show interface** *interface* **switchport**. The output shows the administrative and operational modes as well as the VLAN that the port is currently assigned to.

Example 1.4 shows that a switch port was administratively configured as an access port and is currently operational as an access port. This port is operational in VLAN 10 (Finance).

EXAMPLE 1.4 **Verifying Interface Mode by Using the switchport Command**

```
SW1#
SW1# show interface GigabitEthernet 0/1 switchport
Name: Gi0/1
Switchport: Enabled
Administrative Mode: static access
Operational Mode: static access
Administrative Trunking Encapsulation: negotiate
Operational Trunking Encapsulation: native
Negotiation of Trunking: Off
Access Mode VLAN: 10 (Finance)
<... output omitted ...>
```

# 802.1Q Trunking

A switch port normally carries traffic for a particular VLAN. A trunk link is a special type of connection that carries traffic for multiple VLANs. Unlike an access port, which carries the traffic for only one VLAN, a trunk link typically interconnects switches and carries traffic for multiple VLANs simultaneously. An IEEE 802.1Q trunk is a point-to-point link between two devices that is capable of carrying the traffic for multiple VLANs. All the users who are part of a given VLAN carried over a trunk are in the same broadcast domain.

An 802.1Q trunk provides VLAN identification by adding a 4-byte tag to the Ethernet frame as it leaves the trunk port. Because the frame has been changed, a new frame check sequence (FCS) must be computed and added to the frame. Once the Ethernet frame reaches the receiving switch, the 802.1Q encapsulation header tells the receiving switch what VLAN that frame belongs to. By default, on Cisco Catalyst switches, all configured VLANs are carried over a trunk link unless you specifically remove them.

Because Ethernet is shared and more than two devices could be connected to this medium, all devices must still communicate even if they do not understand 802.1Q tagging. For this reason, 802.1Q also defines a native VLAN. There is one VLAN that is untagged (VLAN 1 by default). All the other VLANs are tagged with a VLAN identifier (VID). Both ends of the trunk link must agree on the native VLAN. Native VLAN mismatch on either side of the 802.1Q trunk link could potentially cause a Layer 2 loop. Cisco Catalyst switches use Cisco Discovery Protocol (CDP) to warn about native VLAN mismatches.

You can statically configure a switch port as a trunk by using the interface-level command **switchport mode trunk**. You can also specify the native VLAN that will be used by using the command **switchport trunk native vlan** *vlan-id*. This command needs to match at both ends of the trunk link to prevent a native VLAN mismatch. You can also specify the VLANs allowed on the trunk by using the following keywords:

- **vlan:** Specifies an explicit list of VLANs, separated by commas or dashes.
- **all:** Specifies all active VLANs (1 through 4094).
- **add:** Indicates a list of VLANs to add to an already configured allowed list.
- **except:** Indicates that all VLANs (1 through 4094) will be added except for this specified list of VLANs.
- **remove:** Specifies a list of VLANs that will be removed from an already configured list.

On some Catalyst switch models, you might need to manually configure the encapsulation protocol before enabling trunking on the port. You do this by using the command **switchport trunk encapsulation dot1q**.

Example 1.5 shows the configuration of an interface in trunk mode. It also demonstrates the configuration of VLAN 100 as the native VLAN on the trunk. Finally, the example demonstrates that VLANs 10 and 20 are being explicitly allowed on the trunk.

EXAMPLE 1.5 **802.1Q Trunk Configuration**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# interface GigabitEthernet 0/3
SW1(config-if)# switchport mode trunk
SW1(config-if)# switchport trunk native vlan 100
SW1(config-if)# switchport trunk allowed vlan 10,20
SW1(config-if)# end
SW1#
```

For verification, the command **show interface trunk** lists all the interfaces on the switch that are configured and operating in trunk mode in Example 1.6. The output also shows the encapsulation type (802.1Q), the native VLAN (100), and the VLANs allowed on the trunk (VLANs 10 and 20).

In addition, Example 1.6 shows the output of the **show interface** *interface* **switchport** command. The output shows the administrative and operational modes of the port as trunk. The output also shows the administrative and operational encapsulation type as being 802.1Q.

EXAMPLE 1.6 **802.1Q Trunk Verification**

```
SW1# show interface trunk

Port        Mode             Encapsulation  Status        Native vlan
Gi0/3       on               802.1q         trunking      100

Port        Vlans allowed on trunk
Gi0/3       10,20

Port        Vlans allowed and active in management domain
Gi0/3       10,20

Port        Vlans in spanning tree forwarding state and not pruned
Gi0/3       10,20
```

```
SW1#
SW1# show interface GigabitEthernet 0/3 switchport
Name: Gi0/3
Switchport: Enabled
Administrative Mode: trunk
Operational Mode: trunk
Administrative Trunking Encapsulation: dot1q
Operational Trunking Encapsulation: dot1q
<... output omitted ...>
```

# Dynamic Trunking Protocol (DTP)

DTP is a Cisco-proprietary protocol that is used to dynamically negotiate the formation of a trunk. DTP is enabled by default, but it can be disabled. DTP manages trunk negotiation only if the port on the neighbor switch is configured in a trunk mode that supports DTP.

**ExamAlert**

Before you take the ENCOR exam, make sure you are familiar with the following modes for the formation of trunk links.

The most common way of setting up trunk links between two devices to pass VLANs is to statically set the switch ports as trunks. However, interfaces on Cisco Catalyst switches support different trunking modes, with the help of DTP for formation of trunk links. These can be configured as follows:

- **dynamic auto:** The interface is able to convert the link to a trunk link. However, this interface can become a trunk link if the other end is set to trunk or desirable mode. In this mode, the interface acts as an access port but listens for DTP packets. It responds to DTP packets and, upon successful negotiation, becomes a trunk link. The command **switchport mode dynamic auto** is used to place an interface in this mode.
- **dynamic desirable:** The interface actively tries to convert the link to a trunk link. The interface becomes a trunk interface if the other end is set to trunk, desirable, or auto mode. In these modes, the interface acts as an access port but listens for and advertises DTP packets to the other end of the link to establish a trunk link. The command **switchport mode dynamic desirable** is used to configure the interface in this mode.

- **trunk:** The interface is placed in permanent trunk mode and negotiates to convert the neighboring link into a trunk link. The command **switchport mode trunk** is used to place a switch port in this mode.

As shown in Table 1.1, a combination of DTP modes can produce either an access port or a trunk port. Table 1.1 illustrates the DTP configuration results on both ends of a link connected to Cisco Catalyst switches.

TABLE 1.1 **DTP Configuration Results**

| | Dynamic Auto | Dynamic Desirable | Trunk | Access |
|---|---|---|---|---|
| **Dynamic Auto** | Access | Trunk | Trunk | Access |
| **Dynamic Desirable** | Trunk | Trunk | Trunk | Access |
| **Trunk** | Trunk | Trunk | Trunk | Limited Connectivity |
| **Access** | Access | Access | Limited Connectivity | Access |

You can use the **switchport nonegotiate** command to prevent an interface from generating DTP frames. This command can only be used when the interface switch port mode is trunk. To establish a trunk link, the interface at the other end needs to be manually configured as a trunk.

Example 1.7 shows the configuration of a trunk link between SW1 and SW3. This example shows how to configure both switches to actively negotiate the formation of a trunk by using the **switchport mode dynamic desirable** command.

EXAMPLE 1.7 **Configuring an 802.1Q Trunk by Using DTP**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# interface GigabitEthernet 0/0
SW1(config-if)# switchport mode dynamic desirable
SW1(config-if)# end

SW3#
SW3# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW3(config)# interface GigabitEthernet 0/0
```

```
SW3(config-if)# switchport mode dynamic desirable
SW3(config-if)# end
SW3#
```

The output of the **show interface trunk** command in Example 1.8 shows that the trunk was formed using the dynamic desirable mode, and the port status is trunking. The **show interface** *interface* **switchport** command shows that the port administrative mode is dynamic desirable but the operational mode is trunk.

EXAMPLE 1.8 **Verifying an 802.1Q Trunk for DTP**

```
SW1#
SW1# show interface trunk

Port        Mode             Encapsulation  Status        Native vlan
Gi0/0       desirable        802.1q         trunking      1
Gi0/3       on               802.1q         trunking      100

Port        Vlans allowed on trunk
Gi0/0       1-4094
Gi0/3       10,20

Port        Vlans allowed and active in management domain
Gi0/0       1,10,20,200
Gi0/3       10,20

Port        Vlans in spanning tree forwarding state and not pruned
Gi0/0       none
Gi0/3       10,20
SW1#
SW1# show interface GigabitEthernet 0/0 switchport
Name: Gi0/0
Switchport: Enabled
Administrative Mode: dynamic desirable
Operational Mode: trunk
Administrative Trunking Encapsulation: dot1q
Operational Trunking Encapsulation: dot1q
Negotiation of Trunking: On
<... output omitted ...>
```

# VLAN Trunking Protocol (VTP)

VTP is a Cisco-proprietary protocol that reduces the burden of provisioning VLANs on switches. It is a Layer 2 messaging protocol that maintains VLAN

configuration consistency by managing the addition, deletion, and renaming of VLANs in the Layer 2 portion of the network. With modern enterprise campus architectures, you often route between the core and distribution layers—and even down to the access layer. As a result, the use of Layer 2 or VTP to propagate VLAN changes is limited to the areas in the network where you trunk traffic. VTP minimizes misconfigurations and configuration inconsistencies that can cause problems, such as duplicate VLAN names, incorrect VLAN-type specifications, and security violations.

Before you can create VLANs, you need to decide whether to use VTP in the network environment. When you use VTP, you can make configuration changes centrally on one or more switches and have those changes automatically communicated to the other switches within the portion of the network where you have trunk links. Without VTP, you cannot send information about VLANs to other switches; the VLANs would need to be manually created. Manual creation of VLANs on every switch in an environment may not be an issue if you are dealing with a small number of VLANs. However, it may become an administrative burden when dealing with a large number of VLANs and switches.

VTP is designed to work in an environment where updates are made on a single switch and are sent through VTP messages to other switches in the domain. It does not work well in a situation where multiple updates to the VLAN database occur simultaneously on switches in the same domain; this situation would result in an inconsistency in the VLAN database. VTP uses configuration revision numbers, which increment as you make changes in the VTP database. There is a risk of deleting an entire VLAN database if you introduce a switch with a higher configuration revision number into your production environment. This is one of the major disadvantages of using VTP.

Before looking at a configuration example of VTP, let's take a brief look at some of the terminology used in a VTP configuration:

- **VTP domain:** A VTP domain (also called a VLAN management domain) consists of one device or several interconnected devices under the same administrative responsibility sharing the same VTP domain name. A device can be in only one VTP domain, and the domain name is case-sensitive. You make global VLAN configuration changes for the domain.

ExamAlert

For the ENCOR exam, make sure you are familiar with the following VTP modes and how they behave in Layer 2 networks.

- **VTP mode:** A switch participating in VTP can be in one of four modes:
  - **Server:** In VTP server mode, the device can create, modify, and delete VLANs and specify other configuration parameters (such as the VTP version) for the entire VTP domain. VTP servers advertise their VLAN configurations to other devices in the same VTP domain. This is the default mode. In VTP server mode, configurations are saved in the vlan.dat file in flash.
  - **Client:** In VTP client mode, a switch functions like a VTP server and transmits and receives VTP updates on its trunks, but you cannot create, change, or delete VLANs on a VTP client. VLANs are configured on another device in the domain that is in VTP server mode. In VTP Versions 1 and 2 in VTP client mode, VLAN configurations are saved in vlan.dat in flash. In VTP Version 3, VLAN configurations are saved in NVRAM in client mode.
  - **Transparent:** In VTP transparent mode, a switch does not participate in VTP. A VTP transparent device does not advertise its VLAN configuration and does not synchronize its VLAN configuration based on received advertisements. However, in VTP transparent mode, devices do forward VTP advertisements that they receive from other devices through their trunk interfaces. In transparent mode, you can create, modify, and delete VLANs on a device. When a device is in VTP transparent mode, the VTP and VLAN configurations are saved in NVRAM. In this mode, the VTP mode and domain name are saved in the device running configuration.
  - **Off:** A switch in VTP off mode functions in the same manner as a VTP transparent device, except that it does not forward any received VTP advertisements from servers or clients on trunks.
- **VTP advertisements:** Each device in a VTP domain sends periodic global configuration advertisements from each trunk port to the reserved multicast address 01-00-0C-CC-CC-CC. Neighboring devices receive these advertisements and update their VTP and VLAN configurations as necessary. The following global domain information is distributed via VTP advertisements:
  - VTP domain name
  - VTP configuration revision number
  - Update identity and update timestamp
  - MD5 digest VLAN configuration, including maximum transmission unit (MTU) size for each VLAN
  - Frame format

- The following VLAN information for each configured VLAN is distributed via VTP advertisements:
  - VLAN IDs
  - VLAN name
  - VLAN type
  - VLAN state
  - Additional VLAN configuration information specific to the VLAN type
  - VTP Version 3 advertisements also include the primary server ID, an instance number, and a start index.
- **VTP versions:** You can implement three different versions of VTP on Cisco Catalyst switches. The same version needs to be configured on all the switches in the same domain. VTP Version 1 is the default on Cisco Catalyst switches. VTP Versions 1 and 2 support propagation of normal-range VLANs (VLANs 1 through 1005). VTP Version 3 supports propagation of the full range of VLANs from 1 through 4094. Extended range VLANs (VLANs 1006 through 4094) are supported only in VTP Version 3. You cannot convert from VTP Version 3 to VTP Version 2 if extended VLANs are already configured in the domain.

VTP supports multiple VTP servers in a domain, and these servers process VTP updates from other VTP servers just as a client does. However, with VTP Version 3, one VTP server must be set as the primary. A VTP primary server updates the database information and sends updates that are honored by all devices in the system. A VTP secondary server can only back up the updated VTP configurations received from the primary server to its NVRAM. VTP Version 3 can also be used to propagate Multiple Spanning Tree (MST) protocol database information. (MST is covered later in this chapter.)

Example 1.9 shows a basic VTP configuration. The VTP domain is defined as ExamCram using the command **vtp domain** *domain-name*, the VTP mode is set to server using the command **vtp mode** *mode*, and the VTP password is configured as Cisco123 with the command **vtp password** *password*. In VTP Version 3, you configure the VTP server as primary by using the **vtp primary** command.

EXAMPLE 1.9 **VTP Configuration**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
```

```
SW1(config)# vtp domain ExamCram
Changing VTP domain name from NULL to ExamCram
SW1(config)# vtp mode server
Device mode already VTP Server for VLANS.
SW1(config)# vtp password Cisco123
Setting device VTP password to Cisco123
SW1(config)# end
SW1#
SW1# vtp primary
This system is becoming primary server for feature vlan
No conflicting VTP3 devices found.
Do you want to continue? [confirm]
SW1#
```

Once you have the trunk link configured to the other switches in the environment and have the same VTP configuration, you should see the VLANs being propagated.

Example 1.10 shows the verification of the VTP status. The command **show vtp status** shows the current VTP version, the defined VTP domain mode, the VTP operating mode, the number of existing VLANs, the configuration revision number, and the primary ID.

EXAMPLE 1.10 **Verifying VTP**

```
SW3#
SW3# show vtp status
VTP Version capable             : 1 to 3
VTP version running             : 3
VTP Domain Name                 : ExamCram
VTP Pruning Mode                : Disabled
VTP Traps Generation            : Disabled
Device ID                       : 5254.000b.8000

Feature VLAN:
--------------
VTP Operating Mode              : Server
Number of existing VLANs        : 9
Number of existing extended VLANs : 0
Maximum VLANs supported locally   : 4096
Configuration Revision          : 2
Primary ID                      : 5254.001a.8000
Primary Description             : SW1
MD5 digest                      : 0x84 0x47 0x3E 0x7A 0x46 0xAB 0x31
                                  0x22
                                  0x65 0x06 0x36 0x19 0x26 0xF3 0xBA
                                  0x35
<... output omitted ...>
```

# Inter-VLAN Routing

Recall that in a switched network, VLANs separate devices into different broadcast domains and Layer 3 subnets. Devices within a VLAN can communicate with one another without the need for routing. However, devices in separate VLANs need a routing device to communicate with devices in other VLANs.

For routing, Layer 2-only switches require a Layer 3 device. Traditionally, this is a router's function. A router needs to have a logical or physical connection to each VLAN to forward packets between them. The process of moving packets from one VLAN to another is known as *inter-VLAN routing*.

Inter-VLAN routing can be performed by a router that is external to the Layer 2 switching environment. This router would have a connection to each of the VLANs. However, such a design is not scalable. Another option is to get an external router connected to a switch by using a trunk link that carries all of the VLANs; this configuration is referred to as a *router on a stick*. The router's physical interface is divided into multiple logical, addressable interfaces with one per VLAN. Each VLAN is associated with a subinterface via the **encapsulation dot1q** *vlan-id* command. The last and most commonly used option is to use a Layer 3, or multilayer, switch.

A Layer 3 switch incorporates the routing capability within the switch. When a switch receives a packet and determines that the packet belongs to another VLAN, it sends the packet to the appropriate port on the other VLAN. A VLAN interface or SVI is a Layer 3 interface created to provide communication between the VLANs as well as other subnets in the network. To route traffic between VLANs, you must create and configure a VLAN interface for each VLAN. Each VLAN requires only one VLAN interface.

Example 1.11 shows the basic configuration of SVIs and their verification. You create an SVI when you enter the interface configuration mode for a VLAN by using the **interface vlan** *vlan-id* command. Example 1.11 shows how to create two SVIs and assign an IP address for each of them. For verification, this example shows the output of the **show ip interface brief** command and excludes the unassigned interface to show the created SVIs and their status.

EXAMPLE 1.11 **Configuring and Verifying SVIs**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# interface vlan 10
SW1(config-if)# ip address 10.10.10.1 255.255.255.0
```

```
SW1(config-if)# no shutdown
SW1(config)# interface vlan 20
SW1(config-if)# ip address 10.10.20.1 255.255.255.0
SW1(config-if)# no shutdown
SW1(config-if)# end
SW1#

SW1# show ip interface brief | exclude unassigned
Interface              IP-Address      OK? Method Status      Protocol
Vlan10                 10.10.10.1      YES manual up                up
Vlan20                 10.10.20.1      YES manual up                up

SW1#
```

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which command is used to assign a port to VLAN 10?
   - ❍ **A. switchport mode access**
   - ❍ **B. switchport access vlan 10**
   - ❍ **C. switchport mode access**
   - ❍ **D. switchport access 10 vlan**

2. What type of port carries frames for multiple VLANs by tagging and untagging the frames?
   - ❍ **A.** EtherChannel port
   - ❍ **B.** Access port
   - ❍ **C.** SVI
   - ❍ **D.** Trunk port

3. Which command is used to prevent an interface from generating Dynamic Trunking Protocol (DTP) frames?
   - ❍ **A. switchport dtp disable**
   - ❍ **B. switchport negotiate**
   - ❍ **C. switchport nonegotiate**
   - ❍ **D. no switchport negotiate**

4. Which of the following commands on a switch port allows the port to actively try to convert to a trunk link?
   - ❍ **A. switchport mode dynamic desirable**
   - ❍ **B. switchport mode dynamic auto**
   - ❍ **C. switchport mode access**
   - ❍ **D. switchport nonegotiate**

# Answers

1. **B** is correct. In this case, the **switchport access vlan 10** command would be used to assign the switch port to VLAN 10.
2. **D** is correct. An IEEE 802.1Q trunk port is a point-to-point link between two devices that is capable of carrying the traffic for multiple VLANs.
3. **C** is correct. The **switchport nonegotiate** command prevents trunk ports from generating DTP frames to dynamically form trunk links.
4. **A** is correct. With the **switchport mode dynamic desirable** command, the interface actively tries to convert the link to a trunk link. The interface becomes a trunk interface if the other end is set to trunk, desirable, or auto mode.

# Spanning Tree Protocol Overview

This section explores the configuration and verification of Cisco Rapid Per VLAN Spanning Tree (RPVST+), Multiple Spanning Tree Protocol (MST), and the various Spanning Tree Protocol tuning and protection mechanisms.

Spanning Tree Protocol prevents loops from being formed in a network when switches are interconnected via multiple paths. Spanning Tree Protocol implements the IEEE 802.1D algorithm by exchanging bridge protocol data unit (BPDU) messages with other switches to detect loops and then removes the loops by shutting down selected switch interfaces. This algorithm guarantees that there is one and only one active path between two network devices. By default, a single instance of Spanning Tree Protocol runs on each configured VLAN—assuming that you do not manually disable Spanning Tree Protocol. Spanning Tree Protocol can be enabled and disabled on a per-VLAN basis.

In general, switches send and receive spanning-tree frames at regular intervals. The switches do not forward these frames but use the frames to construct loop-free paths. Multiple active paths between end stations cause loops in a network. If a loop exists in a network, end stations might receive duplicate messages, and switches might learn end station MAC addresses on multiple Layer 2 interfaces. Learning the MAC address for a device on multiple interfaces results in an unstable network.

The following factors contribute to a spanning-tree network topology being stable and active:

- The unique bridge ID associated with each VLAN on each bridge
- The spanning-tree path cost to the spanning-tree root
- The port identifier (port priority and MAC address) associated with each Layer 2 interface

Switches use BPDUs to exchange Spanning Tree Protocol information, specifically for root switch election and for loop identification. By default, BPDUs are sent every 2 seconds, and they are of three types:

- **Configuration BPDUs:** Used to identify the root bridge, root ports, designated ports, and blocking ports.
- **Topology Change Notification (TCN) BPDUs:** Used to forward topology changes on the root port toward the root bridge. A topology change can result from a link failure, a switch failure, or a port transitioning to the forwarding state.

- **Topology Change Acknowledgement (TCA) BPDUs:** Used by an upstream switch to acknowledge a TCN BPDU.

To summarize, a switch continually receives configuration BPDUs from the root switch on its root port in normal switch operation. When there is a change in the topology, the switch sends a TCN BPDU on its root port.

Spanning Tree Protocol defines a tree with a root switch and a loop-free path from the root to all switches in the Layer 2 network. Spanning Tree Protocol forces redundant data paths into a standby (blocked) state. If a network segment in the spanning tree fails and a redundant path exists, the spanning-tree algorithm recalculates the spanning-tree topology and activates the standby path. When two ports on a switch are part of a loop, the spanning-tree port priority and port path cost setting determine which port is put in the forwarding state and which port is put in the blocking state. The spanning-tree port priority value represents the location of an interface in the network topology and how well located it is to pass traffic. The spanning-tree port path cost value represents the media speed.

> **Note**
>
> The terms *bridge* and *switch* are used interchangeably for the rest of this chapter. Next, we will look at the root bridge election, root port election, and designated port election in more detail.

# Root Bridge, Root Port, and Designated Port Elections

For each VLAN, the switch with the highest bridge ID (that is, the lowest numerical ID value) is elected as the root bridge. This is true for Cisco's PVST+ and RPVST+ (which is Cisco's enhanced implementation of the standards-based RSTP). If all switches are configured with the default priority (32768), the switch with the lowest MAC address for each VLAN becomes the root bridge. There can be only one root bridge per VLAN in a network. The bridge priority value occupies the most significant bits of the bridge ID. When the bridge priority value is changed, you change the probability that the switch will be elected as the root bridge. Configuring a lower value increases the probability; configuring a higher value decreases the probability.

The Spanning Tree Protocol root bridge is the logical center of each spanning-tree topology in a network. All paths that are not needed to reach the root

bridge from anywhere in the network are placed in Spanning Tree Protocol blocking mode.

BPDUs contain information about the transmitting bridge and its ports, including bridge and MAC addresses, bridge priority, port priority, and path cost. Spanning Tree Protocol uses this information to elect the root bridge for the Spanning Tree Protocol instance, elect the root port leading to the root bridge, and determine the designated port for each segment.

In Example 1.12, notice that the **show spanning-tree** command explicitly shows that this is the root bridge for VLAN 1. The SW1 MAC address in this case is also the same as the root bridge, which indicates that this is indeed the root bridge for VLAN 1. Also, notice that all of the ports for VLAN 1 on this switch are in designated state, which is always the case for the root bridge.

EXAMPLE 1.12 **Verifying Root Bridge ID and Port State**

```
SW1#
SW1# show spanning-tree

VLAN0001
  Spanning tree enabled protocol ieee
  Root ID     Priority    24577
              Address     5254.001a.37c2
              This bridge is the root
              Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec

  Bridge ID   Priority    24577  (priority 24576 sys-id-ext 1)
              Address     5254.001a.37c2
              Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec
              Aging Time  15  sec

Interface           Role Sts Cost      Prio.Nbr Type
------------------- ---- --- --------- -------- --------------------
Gi0/0               Desg FWD 4         128.1    P2p
Gi0/1               Desg FWD 4         128.2    P2p
Gi0/2               Desg FWD 4         128.3    P2p
```

The root port provides the best path (lowest cost) when a switch forwards packets to the root bridge. After the root bridge election, each non-root bridge must figure out where it is in relationship to the root bridge. As it relates to root port election, the path cost value is used. The path cost is the cumulative cost of all links toward the root bridge. The root port, in this case, is the port with the lowest cumulative cost. However, if two ports have the same cost, the sender's bridge ID (BID) is used to break the tie first, and if they are the same, then the sender's port ID is used to break the tie.

Spanning Tree Protocol cost is based on the bandwidth of the link. This value can be manually changed, although manually changing this value is not common. The higher the bandwidth of a link, the lower the cost. Cisco Catalyst switches support two methods, which can be changed by using the **spanning-tree pathcost method** *method* command:

- **Short method:** In the short method, 16-bit (short) default port costs are assigned to each port using a formula that is based on the port bandwidth. You can manually assign port costs between 1 and 65535.
- **Long method:** In the long method, 32-bit (long) default port cost values are assigned to each port using a formula that is based on the port bandwidth. You can also manually assign port costs between 1 and 200,000,000. The formula for obtaining default 32-bit port costs is to divide the bandwidth of the port by 200,000,000.

Table 1.2 shows the default port cost values using the short method.

TABLE 1.2 **Default Port Cost Values Using the Short Method**

| Port Speed | Default Cost Value | Default Range |
|---|---|---|
| 10 Mbps | 100 | 1 to 65,535 |
| 100 Mbps | 19 | 1 to 65,535 |
| 1 Gbps | 4 | 1 to 65,535 |

Table 1.3 shows the default port cost values using the long method.

TABLE 1.3 **Default Port Cost Values Using the Long Method**

| Port Speed | Recommended Value | Available Range |
|---|---|---|
| 1 Mbps | 20,000,000 | 1 to 200,000,000 |
| 10 Mbps | 2,000,000 | 1 to 200,000,000 |
| 100 Mbps | 200,000 | 1 to 200,000,000 |
| 1 Gbps | 20000 | 1 to 200,000,000 |
| 10 Gbps | 2000 | 1 to 200,000,000 |

Example 1.13 shows how to change the path cost method by using the **spanning-tree pathcost method long** command. This change is verified by using the **show spanning-tree vlan 1** and **show spanning-tree summary** commands. Note that the interface's cost is now 20000 to account for 1 Gbps interfaces and also the output which indicates that the long path cost method is now being used.

EXAMPLE 1.13 **Changing and Verifying the Spanning Tree Protocol Path Cost Method**

```
SW3#
SW3# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW3(config)# spanning-tree pathcost method long
SW3(config)# end
SW3#
SW3# show spanning-tree vlan 1

VLAN0001
  Spanning tree enabled protocol ieee
  Root ID     Priority    24577
              Address     5254.000b.9f89
              This bridge is the root
              Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec

  Bridge ID   Priority    24577  (priority 24576 sys-id-ext 1)
              Address     5254.000b.9f89
              Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec
              Aging Time  300 sec

Interface           Role Sts Cost      Prio.Nbr Type
------------------- ---- --- --------- -------- ----------------------
Gi0/0               Desg FWD 20000     128.1    P2p
Gi0/1               Desg FWD 20000     128.2    P2p
Gi0/2               Desg FWD 20000     128.3    P2p
Gi0/3               Desg FWD 20000     128.4    P2p

SW3#
SW3# show spanning-tree summary
<... output omitted ...>
Configured Pathcost method used is long
<... output omitted ...>
SW3#
```

There is one designated port per segment. The designated port is selected on the bridge with the lowest-cost path to the root bridge and is responsible for forwarding traffic on that segment. If there are multiple paths with equal cost to the root bridge, Spanning Tree Protocol uses the following criteria to determine the designated and non-designated ports on the segment:

- Lowest path cost to the root bridge
- Lowest sender bridge ID
- Lowest sender port ID

All ports that are not root or designated ports are non-designated ports. A non-designated port goes into blocking state to prevent loops in the Spanning Tree Protocol topology.

To participate in Spanning Tree Protocol, a switch port goes through different states. A switch port starts by being in the disabled state; after it is enabled, it moves through several states before it reaches the forwarding state, provided that the port is not a designated port or a root port. If the port is not in any of these states, it is moved into the blocking state.

The following list briefly explains the various states a Layer 2 port could be in when participating in 802.1D Spanning Tree Protocol:

- **Disabled:** The Layer 2 interface does not participate in Spanning Tree Protocol and is not forwarding frames because the port is administratively shut down.
- **Blocking:** The Layer 2 interface does not participate in frame forwarding. In this state, the port does not send or receive data but receives BPDUs from other switches. The switch does not modify the MAC address table in this state.
- **Listening:** This is the first transitional state after the blocking state, when Spanning Tree Protocol determines that the Layer 2 interface must participate in frame forwarding. A port in the listening state cannot send or receive frames, but it is allowed to send and receive BPDUs.
- **Learning:** The Layer 2 interface prepares to participate in frame forwarding. After the listening state expires (15 seconds by default), the port transitions to learning state. In this state, the port sends and receives BPDUs and can learn and add new MAC addresses to the MAC address table. In this state, the port is not yet sending data frames.
- **Forwarding:** The Layer 2 interface forwards frames. After the learning state expires (15 seconds by default), the port moves into the forwarding state. In the forwarding state, the port sends and receives frames and sends and receives BPDUs.

You can tune three timers when configuring Spanning Tree Protocol. These can be tuned globally for all VLANs or by specifying the timers for a particular VLAN. The timers deal with the following times:

- **Hello time:** The hello time is the time that a BPDU is advertised out a port. It can be a value within the range 1 to 10 seconds, with the default

hello time being 2 seconds. It can be configured with the **spanning-tree vlan** *vlan-id* **hello-time** *hello-time* command.

- **Forward delay time:** The forward delay time is the time that a port stays in the listening and learning state. It can be a value within the range 15 to 30 seconds, with the default forward delay time being 15 seconds. It can be configured with the **spanning-tree vlan** *vlan-id* **forward-time** *forward-time* command.
- **Max age time:** The max age time is the maximum length of time that passes before a bridge saves its configuration BPDU information. It can be a value within the range 6 to 40 seconds, with the default max age time being 20 seconds. It can be configured with the **spanning-tree vlan** *vlan-id* **max-age** *max-age* command.

The timers are set at the root bridge level. Although the timers can be modified on other switches that are not the root bridge, such settings have no effect as the advertised timers from the root bridge are used.

# Rapid Spanning Tree Protocol (RSTP)

Rapid Spanning Tree Protocol (RSTP; 802.1w) includes enhancements to the IEEE 802.1D Spanning Tree Protocol standard. The 802.1D Spanning Tree Protocol standard was designed at a time when the recovery of connectivity after an outage within a minute or so was considered adequate performance. With the advent of Layer 3 switching in LAN environments, bridging now competes with routed solutions where protocols such as OSPF and EIGRP can provide an alternate path in less time.

Cisco enhanced the original IEEE 802.1D Spanning Tree Protocol specification with features such as Uplink Fast, Backbone Fast, and Port Fast to speed up the convergence time in a switched network. The drawback of these mechanisms is that they are proprietary and need additional configuration to be implemented in a switched network.

RPVST+ is Cisco's enhancement of the standards-based RSTP (802.1w), which uses Cisco's PVST+ and provides a separate instance of RSTP for each VLAN.

Example 1.14 shows the configuration and verification of RPVST+. In this example, the spanning-tree mode is changed using the **spanning-tree mode rapid-pvst** command. Verification is done using the **show spanning-tree** command.

EXAMPLE 1.14 **Configuring and Verifying RPVST+**

```
SW3#
SW3# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW3(config)# spanning-tree mode rapid-pvst
SW3(config)# end
SW3#
SW3# show spanning-tree

VLAN0001
  Spanning tree enabled protocol rstp
  <... output omitted ...>
```

With RSTP, the 802.1D terminology and parameters remain primarily the same. In most cases, RSTP performs better than proprietary extensions of Cisco without any additional configuration. 802.1w can also revert to 802.1D in order to interoperate with legacy switches on a per-port basis. There are only three port states left in RSTP that correspond to the three possible operational states. However, the 802.1D disabled, blocking, and listening states are merged into a unique 802.1w discarding state.

Table 1.4 shows the different port states in Spanning Tree Protocol (802.1D) and RSTP (RSTP 802.1w).

TABLE 1.4 **Port States in Spanning Tree Protocol (802.1D) and RSTP (RSTP 802.1w)**

| Spanning Tree Protocol 802.1D Port State | RSTP 802.1w Port State | Is Port Included in Active Topology? | Is Port Learning MAC Addresses? |
|---|---|---|---|
| Disabled | Discarding | No | No |
| Blocking | Discarding | No | No |
| Listening | Discarding | Yes | No |
| Learning | Learning | Yes | Yes |
| Forwarding | Forwarding | Yes | Yes |

There are a number of port roles in RSTP. The spanning-tree algorithm determines a port's role based on BPDUs. The following port roles are possible:

- **Root port:** The root port is the port that receives the best BPDU. It is the port that is the closest to the root switch in terms of path cost. The spanning-tree algorithm elects a single root switch in the whole network

for each VLAN. The root switch is the only switch in the network that does not have a root port.

- **Designated port:** The designated port is the port that can send the best BPDU on the segment to which it is connected. On a given segment, there can be only one path toward the root switch. All switches connected to a given segment listen to the BPDUs of each segment and agree on the switch that sends the best BPDU as the designated switch for the segment. The corresponding port on the switch is the designated port for that segment.
- **Alternate port:** The alternate port is a switch port that provides an alternate path toward the root switch through a different switch. The alternate port makes a transition to a designated port if the current designated path fails.
- **Backup port:** The backup port is the port that provides link redundancy toward the root switch. It acts as a backup for the path and is provided by a designated port toward the leaves of the spanning tree. A backup port exists only if there are multiple links toward the root switch.
- **Disabled port:** A disabled port has no role in the operation of spanning tree.

RSTP provides rapid convergence for edge ports, new root ports, and ports connected through point-to-point links as follows:

- **Edge ports:** When you configure a port as an edge port on an RSTP switch, the edge port immediately transitions to the forwarding state. This immediate transition was previously the Cisco-proprietary feature PortFast. Only the port that connects to a single end station should be configured as an edge port. Edge ports do not generate topology changes when the link changes.
- **Root ports:** If RSTP selects a new root port, it blocks the old root port and immediately transitions the new root port to the forwarding state.
- **Point-to-point links:** If you connect a port to another port through a point-to-point link and the local port becomes a designated port, it negotiates a rapid transition with the other port by using the proposal-agreement handshake to ensure a loop-free topology.

After looking at the various Spanning Tree Protocol tuning and protection mechanisms in the following section, we cover another type of Spanning Tree

Protocol deployment, Multiple Spanning Tree Protocol (MST). You will use some of these tuning and protection mechanisms next, to tune and protect Spanning Tree Protocol, including MST.

# Spanning Tree Protocol Tuning and Protection Mechanisms

To create a stable Spanning Tree Protocol topology, you need to put in place mechanisms to guarantee the root bridge placement in the network. It is also important to have a potential backup root bridge for the Layer 2 topology. The guaranteed placement of the root bridge is critical to prevent another switch from unintentionally taking over the root bridge role.

This section looks at techniques for guaranteeing the root bridge and secondary root bridge placement in a network by modifying the system priority. This section also looks at other Spanning Tree Protocol protection mechanisms, including Root Guard, PortFast, BPDU Guard, BPDU Filter, Loop Guard, Bridge Assurance, and Unidirectional Link Detection (UDLD). First, let's consider switch priorities.

# Switch Priorities Overview

Ideally, a root bridge is placed at the Layer 2/Layer 3 boundary. The placement of the root bridge is an important decision because you want to minimize the number of hops to the furthest switch in the topology. Also, the root switch needs to have enough capabilities to handle cross-switch traffic. Generally, the root switch is at the distribution layer or at the core/distribution layer in a collapsed core design, and you can specify another switch as the secondary root bridge. In most cases, this is done on a per-VLAN basis. To guarantee consistent placement of the root bridge, you can follow these steps:

1. Lower the system priority of the root bridge for each VLAN to the lowest possible value.
2. Lower the system priority of the secondary root bridge for each VLAN to a value slightly higher than the value of the root bridge.
3. Optionally, increase the system priority on all other switches for all VLANs (to guarantee that these switches do not inadvertently become the root bridge at some point in the future).

ExamAlert

For the ENCOR exam, it is important to know the following methods of adjusting the root bridge priority.

You have two options for setting the root bridge's priority and ensuring its consistent placement. The priority can be set using either of the following commands:

- **spanning-tree vlan** *vlan-id* **priority** *priority***:** The priority value can be between 0 and 61440, in increments of 4096.
- **spanning-tree vlan** *vlan-id* **root** {**primary** | **secondary**}**:** This command executes a script to set the root bridge priority. The **primary** keyword sets the priority to 24576, and the **secondary** keyword sets the priority to 28672. However, setting the root bridge priority using this command is true only if the current root bridge is using the default value.

Spanning Tree Protocol uses the extended system ID, the switch priority, and the allocated spanning-tree MAC address to make the bridge ID unique for each VLAN. If you configure the priority as 24576 for VLAN 20, the resulting priority will be 24596 (that is, 24576 plus the VLAN ID 20).

Example 1.15 shows how to make SW3 the root bridge for VLAN 1 with the **root primary** keyword.

For verification, you use the **show spanning-tree vlan 1** command. In this output, note that the root ID priority and the bridge IP priority are the same, which means this is the root bridge. The **root primary** keyword sets the priority to 24577 (that is, 24576 plus 1, where the 1 represents the VLAN ID). Another indicator that this is the root bridge is that all the ports connected to other switches are acting in the designated (Desg) role.

EXAMPLE 1.15 **Configuring and Verifying Spanning Tree Protocol Root Bridge Priority**

```
SW3#
SW3# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW3(config)# spanning-tree vlan 1 root primary
SW3(config)# end
SW3#
```

```
SW3# show spanning-tree vlan 1

VLAN0001
  Spanning tree enabled protocol ieee
  Root ID     Priority    24577
              Address     5254.000b.9f89
              This bridge is the root
              Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec

  Bridge ID   Priority    24577  (priority 24576 sys-id-ext 1)
              Address     5254.000b.9f89
              Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec
              Aging Time  15  sec

Interface           Role Sts Cost      Prio.Nbr Type
------------------- ---- --- --------- -------- ----------------------
Gi0/0               Desg FWD 4         128.1    P2p
Gi0/1               Desg FWD 4         128.2    P2p
Gi0/2               Desg FWD 4         128.3    P2p
Gi0/3
```

Example 1.15 shows how to influence the root bridge placement by using the **root primary** command. You can also manually specify the priority for a particular VLAN, as shown in Example 1.16. Changing the bridge priority to the value 4096 for VLAN 1 would yield an effective priority of 4097 (4096 for the bridge priority plus 1 for the VLAN ID).

EXAMPLE 1.16 **Configuring and Verifying Root Bridge Priority**

```
SW3#
SW3# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW3(config)# spanning-tree vlan 1 priority 4096
SW3(config)# end
SW3#
SW3# show spanning-tree vlan 1

VLAN0001
  Spanning tree enabled protocol rstp
  Root ID     Priority    4097
              Address     5254.000b.9f89
              This bridge is the root
              Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec
```

```
Bridge ID  Priority    4097   (priority 4096 sys-id-ext 1)
           Address     5254.000b.9f89
           Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec
           Aging Time  300 sec
           <... output omitted ...>
```

The optional **diameter** command modifies the Spanning Tree Protocol timers for optimum convergence. The diameter should be configured to reference the maximum number of Layer 2 hops between the root bridge and a switch. The timers do not need to be modified on the non-root switches since the timers are carried in the configuration BPDUs from the root bridge.

Let's now examine the various protection mechanisms that can be configured with Spanning Tree Protocol. Network packets do not decrement the TTL portion of the header as packets are forwarded in a Layer 2 topology. Due to this, a forwarding loop can easily occur when there are multiple active paths between two devices. A forwarding loop with broadcast or multicast traffic can have an even more significant impact as it is forwarded out every switch port, continuing the forwarding loop.

With a forwarding loop, a situation can quickly arise in which there are high CPU utilizations and low free memory space. Because packets are received on different interfaces, the switch needs to move MAC addresses from one interface to the next. Eventually, the user's network applications may slow down, and the switch may crash due to the exhaustion of CPU and memory resources.

To mitigate some of these issues, Cisco added a number of extensions to Spanning Tree Protocol that enhance loop prevention, protect against user misconfiguration, and provide better control over the protocol parameters. The available extensions are Root Guard, PortFast, BPDU Guard, BPDU Filter, Loop Guard, Bridge Assurance, and Unidirectional Link Detection (UDLD). All of these extensions can be used with RPVST+ and MST.

## Root Guard

Root Guard limits the switch ports out of which the root bridge may be negotiated. If a Root Guard–enabled port receives BPDUs that are superior to the BPDUs being sent by the current root bridge, that port will be moved to a root-inconsistent state, which is effectively equal to a Spanning Tree Protocol listening state. Basically, Root Guard prevents downstream switches (which may be rogue or misconfigured) from becoming the root bridge in a topology.

If a superior BPDU is received on a Root Guard–configured port, the Root Guard function places the port in the root-inconsistent state. This prevents the configured designated port with Root Guard from becoming a root port.

Root Guard is configured on a per-interface basis with the **spanning-tree guard root** command, as shown in Example 1.17. It should be configured only on designated ports toward other switches that should never become root bridges.

EXAMPLE 1.17 **Root Guard Configuration**

```
SW3#
SW3# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW3(config)# interface GigabitEthernet 0/0
SW3(config-if)# spanning-tree guard root
SW3(config-if)#
*May  9 17:24:13.486: %SPANTREE-2-ROOTGUARD_CONFIG_CHANGE: Root guard
 enabled on port GigabitEthernet0/0.
SW3(config-if)# end
SW3#
```

# PortFast

Configuring an access port with PortFast causes the port to bypass the listening and learning states and enter the forwarding state immediately. PortFast is typically deployed on Layer 2 access ports that are connected to a single workstation or server. This design allows those devices to connect to the network immediately, without waiting for Spanning Tree Protocol convergence. Interfaces connected to a single workstation or server are not expected to receive BPDUs, and it should be safe to transition these ports to the forwarding state immediately.

Say that you have many end devices, and PortFast is not enabled on access ports. In this case, if end devices are constantly being powered up and powered down, or connected and disconnected, that action will initiate a lot of topology changes in the network. Edge ports do not generate topology changes when the link state changes. PortFast configuration is used to minimize the time that access ports must wait for Spanning Tree Protocol convergence to occur; therefore, it should be used only on access ports. If you enable PortFast on a port connected to a switch, you may inadvertently create a temporary bridging loop.

Spanning Tree Protocol PortFast is enabled on access ports with the interface level command **spanning-tree portfast**. It can also be enabled globally on all access ports with the **spanning-tree portfast default** command. If PortFast is enabled globally and you need to disable it on a specific port, you can use the **spanning-tree portfast disable** command to disable it on that particular interface.

PortFast can also be enabled on a trunk link that goes to a single host. For example, it is common to enable PortFast on a trunk link that goes to a server running a hypervisor. In this case, you want to bring up the trunk link as soon as possible to allow virtual machines (VMs) to communicate across different VLANs. Enabling PortFast on a trunk link that connects to a host, in this case, is done using the **spanning-tree portfast trunk** command. It is important to remember that using this command on interfaces that connect to other switches or bridges can result in bridging loops.

## BPDU Guard

When you configure the PortFast feature, an interface still listens for BPDUs. If a BPDU is received, the port will be moved into a blocking state. However, a loop can only be detected in a finite amount of time; that is, some time is needed to move a port into a blocked state. BPDU Guard protects the integrity of the interface that is configured with Spanning Tree Protocol PortFast. If any BPDU is received on an interface configured with PortFast, that interface is put into an error-disabled state. The port is shut down and must be manually re-enabled or automatically recovered through the error-disabled timeout function.

You can configure BPDU Guard globally on all Spanning Tree Protocol PortFast interfaces by using the **spanning-tree portfast bpduguard default** command. On a per-interface basis, BPDU Guard can be configured with the **spanning-tree bpduguard** {**enable** | **disable**} command. To use error recovery to recover ports that BPDU Guard shut down, you use the command **errdisable recovery cause bpduguard**. The error recovery interval can be set using the **errdisable recovery internal** *time-seconds* command.

Example 1.18 shows the configuration and verification of PortFast and BPDU Guard at both the interface and global configuration levels. For the PortFast verification, note that the port is configured as a spanning-tree PortFast edge. For BPDU Guard verification, note that **spanning-tree bpduguard enable** is in the output.

EXAMPLE 1.18 **Configuring and Verifying PortFast and BPDU Guard**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# interface GigabitEthernet 0/3
SW1(config-if)# spanning-tree portfast
%Warning: portfast should only be enabled on ports connected to a
 single host. Connecting hubs, concentrators, switches, bridges, etc…
 to this interface  when portfast is enabled, can cause temporary
 bridging loops. Use with CAUTION
SW1(config-if)# spanning-tree bpduguard enable
SW1(config-if)# end

SW3#
SW3# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW3(config)# spanning-tree portfast default
%Warning: this command enables portfast by default on all interfaces.
 You should now disable portfast explicitly on switched ports
 leading to hubs,switches and bridges as they may create temporary
 bridging loops.

SW3(config)# spanning-tree portfast bpduguard default
SW3(config)# end
SW3#

SW1# show running-config interface GigabitEthernet 0/3
Building configuration...

Current configuration : 263 bytes
!
interface GigabitEthernet0/3
<... output omitted ...>
 spanning-tree portfast edge
 spanning-tree bpduguard enable
end
SW1#
SW1# show spanning-tree interface GigabitEthernet 0/3 portfast
VLAN0001            enabled
SW1#

SW3#
SW3# show spanning-tree summary
<... output omitted ...>
Portfast Default                         is edge
Portfast Edge BPDU Guard Default         is enabled
<... output omitted ...>
SW3#
```

## BPDU Filter

BPDUs are sent on all ports, even if they are PortFast enabled. In most cases, Spanning Tree Protocol should be enabled in order to prevent loops. However, there are special cases where you need to prevent BPDUs from being sent out; in those cases, you use BPDU Filter. One of those cases is on ports that are configured with PortFast that connects an end station. Typically, you don't want to configure BPDU Filter unless it is absolutely necessary.

One common use case for BPDU Filter is in a service provider environment. BPDU Filter can be configured so that all configuration BPDUs received on a port will be dropped in an environment where a service provider provides Layer 2 Ethernet access for customers. Ideally, the service provider should not share any spanning-tree information with customers because such sharing might jeopardize the stability of the internal spanning-tree topology of the service provider network. By configuring the Spanning Tree Protocol PortFast feature and BPDU Filter on each customer access port, the service provider will not send any configuration BPDUs to customers and will ignore any configuration BPDUs sent from customers.

The BPDU Filter feature can be enabled globally or on a per-interface basis, but the feature operates differently depending on how it is configured. These differences are highlighted next:

- At the global level, you can enable BPDU Filter on PortFast-enabled interfaces by using the **spanning-tree portfast bpdufilter default** global configuration command. This command prevents interfaces that are in a PortFast operational state from sending or receiving any BPDUs. The interfaces still send a few BPDUs at link-up before the switch begins to filter outbound BPDUs. If a BPDU is received on a PortFast-enabled interface, the interface loses its PortFast operational status, BPDU filtering is disabled, and the interface acts as a normal interface.
- At the interface level, you can enable BPDU Filter on any interface by using the **spanning-tree bpdufilter enable** interface configuration command without also enabling the PortFast feature. Enabling BPDU Filter at the interface level prevents the interface from sending or receiving any BPDUs.

As you can see, the main difference in the BPDU Filter implementations is in how it works with the PortFast feature.

## Loop Guard

When one of the ports in a physically redundant topology no longer receives BPDUs (for example, when sending or receiving fails or when filtering is erroneously configured), Spanning Tree Protocol conceives that the topology is loop free. Eventually, the blocking port from the alternate or backup port becomes designated and moves to a forwarding state. This situation creates a loop. You can use Loop Guard to prevent alternate or root ports from becoming designated ports due to a failure that leads to a unidirectional link (that is, one-way traffic on a link or port).

The Loop Guard feature works by performing additional checks. If BPDUs are not received on a non-designated port and Loop Guard is enabled, that port is moved into the Spanning Tree Protocol loop-inconsistent blocking state instead of the listening, learning, and, eventually, forwarding states. Without the Loop Guard feature, the port assumes the designated port role, and the port moves to the Spanning Tree Protocol forwarding state and creates a loop. When BPDU transmission starts again on the interface, the port recovers and begins to transition through the Spanning Tree Protocol states.

The Loop Guard feature is most effective when it is enabled on an entire switched network. Loop Guard can be enabled globally using the **spanning-tree loopguard default** command. On a per-interface basis, Loop Guard can be configured with the **spanning-tree guard loop** command. Loop Guard should not be enabled on PortFast-enabled interfaces because implementing these features together creates a conflict in the root port logic.

Example 1.19 shows an example of the interface level and global configuration of Loop Guard, as well as verification of the configuration.

EXAMPLE 1.19 **Configuring and Verifying Loop Guard**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# interface GigabitEthernet 0/2
SW1(config-if)# spanning-tree guard loop
SW1(config-if)# end
SW1#
```

```
SW3#
SW3# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW3(config)# spanning-tree loopguard default
SW3(config)# end
SW3#

SW3#
SW3# show spanning-tree summary
Switch is in rapid-pvst mode
Root bridge for: VLAN0001, VLAN0010, VLAN0020, VLAN0030, VLAN0200
Extended system ID                        is enabled
Portfast Default                          is edge
Portfast Edge BPDU Guard Default          is enabled
Portfast Edge BPDU Filter Default         is disabled
Loopguard Default                         is enabled
<... output omitted ...>
```

## Bridge Assurance

Bridge Assurance helps prevent looping conditions caused by unidirectional links or a malfunction in a neighboring switch. Bridge Assurance, which is applicable only with RPVST+ and MST, modifies the rules for sending BPDUs. When Bridge Assurance is activated on a port, the port always sends BPDUs, whether it is a root, designated, alternate, or backup port. BPDUs essentially work as a hello mechanism between pairs of interconnected switches. A port that is configured with Bridge Assurance is required to receive BPDUs. If a port does not receive BPDUs, it goes into the blocking state. Thus, both ends of the link must have Bridge Assurance enabled for it to function.

You can think of Bridge Assurance as an extension of the idea that Loop Guard uses. However, whereas Loop Guard is supported on all Cisco switching platforms, Bridge Assurance is not. If network devices support Bridge Assurance, it should be used instead of Loop Guard. However, Loop Guard and Bridge Assurance should not be used at the same time. If the unidirectional problem exists before a link comes up, Loop Guard will not detect such an issue, but Bridge Assurance will.

Bridge Assurance is enabled by default, and you can only disable it globally. Also, Bridge Assurance is enabled only on spanning-tree network ports that are point-to-point links. Bridge Assurance is enabled using the **spanning-tree bridge assurance** command.

Once the link type is set to point-to-point on both ends of the link and Bridge Assurance is set globally, Bridge Assurance starts to send BPDUs out that port, as shown in Example 1.20.

EXAMPLE 1.20 **Bridge Assurance Configuration**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# interface GigabitEthernet 0/0
SW1(config-if)# spanning-tree link-type point-to-point
SW1(config-if)# exit
SW1(config)# spanning-tree bridge assurance
SW1(config)# end
SW1#

SW1# show spanning-tree summary
<... output omitted ...>
Bridge Assurance                        is enabled
<... output omitted ...>
```

# Unidirectional Link Detection (UDLD)

UDLD is a Layer 2 protocol that works with Layer 1 mechanisms to determine the physical status of a link. With UDLD enabled, a switch periodically transmits UDLD packets on the interface. If the packets are not echoed back within a specific time frame, the link is flagged as unidirectional, and the interface is in the error-disabled state. Devices on both ends of the link must support UDLD for the protocol to identify and disable unidirectional links successfully. UDLD does not protect against Spanning Tree Protocol failures caused by software (where the designated switch is not sending BPDUs). Loop Guard can be used in those situations instead.

**ExamAlert**

It would be useful to commit the following UDLD modes to memory for the ENCOR exam.

UDLD has two modes of operation:

- **Normal (default):** In normal mode, UDLD can detect unidirectional links due to misconnected ports on fiber-optic connections. When a

unidirectional link is detected, the port is allowed to continue operations. (UDLD marks the port as having an undetermined state.) A syslog message is also generated.

- **Aggressive:** In aggressive mode, UDLD can also detect unidirectional links due to one-way traffic on fiber-optic and twisted-pair links, as well as misconnected ports on fiber-optic connections. When a unidirectional link is detected, the switch tries to reestablish the link. It sends messages, and if none are returned, it moves the port to an error-disabled state.

UDLD can be enabled globally with the command **udld** {**aggressive** | **enable** | **message time** *interval*}.

Specifying the **aggressive** option puts UDLD in aggressive mode. The **enable** option puts UDLD in normal mode on all fiber-optic ports on the switch. The command **udld** {**enable** | **aggressive** | **disable**} can be used on a per-interface basis, as shown in Example 1.21.

EXAMPLE 1.21 **Configuring and Verifying UDLD**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# udld aggressive
SW1(config)# end
SW1# show udld neighbor
Port     Device Name   Device ID     Port ID    Neighbor State
----     -----------   ---------     -------    --------------
Gi0/0    SW3     1            Gi0/0      Bidirectional
SW1#

SW1# show udld GigabitEthernet 0/0

Interface Gi0/0
---
Port enable administrative configuration setting: Enabled /
  in aggressive mode
Port enable operational state: Enabled / in aggressive mode
Current bidirectional state: Bidirectional
Current operational state: Advertisement - Single neighbor detected
Message interval: 15000 ms
Time out interval: 5000 ms

Port fast-hello configuration setting: Disabled
Port fast-hello interval: 0 ms
Port fast-hello operational state: Disabled
```

```
Neighbor fast-hello configuration setting: Disabled
Neighbor fast-hello interval: Unknown


    Entry 1
    ---
    Expiration time: 33700 ms
    Cache Device index: 1
    Current neighbor state: Bidirectional
    Device ID: SW3
    Port ID: Gi0/0
    Neighbor echo 1 device: 92QV1KDJMV2
    Neighbor echo 1 port: Gi0/0

    TLV Message interval: 15 sec
    No TLV fast-hello interval
    TLV Time out interval: 5
    TLV CDP Device name: SW3
SW1#
```

# Multiple Spanning Tree Protocol (MST)

Multiple Spanning Tree (MST), specified in IEEE 802.1s, enables you to map multiple VLANs to the same spanning-tree instance. This aids in reducing the number of spanning-tree instances needed to support a large number of VLANs. MST also provides for multiple forwarding paths for data traffic and enables load balancing. It improves a network's fault tolerance because a failure in one instance (that is, forwarding path) does not affect other instances (that is, forwarding paths). MST is commonly deployed in the core and distribution layers of a Layer 2 switched network.

Since IEEE 802.1s is the only version of Spanning Tree Protocol that is implemented the same with both Cisco and non-Cisco switches, it is the recommended flavor of Spanning Tree Protocol to run in multi-vendor switch environments. Because of the per-VLAN nature of PVST+ and RPVST+ compared to the standards-based IEEE 802.1d and 802.1w, MST is recommended when interconnecting switches from different vendors.

For switches to participate in MST instances, you must consistently configure the switches with the same MST configuration information. The MST configuration controls the MST region to which each switch belongs. The configuration includes the name of the region, the revision number, and the MST VLAN-to-instance assignment map. You configure a switch for a region by using the **spanning-tree mst configuration** global configuration command,

after which the switch enters the MST configuration mode. In this mode, you can map VLANs to an MST instance by using the **instance** MST configuration command, specify the region name by using the **name** MST configuration command, and set the revision number by using the **revision** MST configuration command.

MST establishes and maintains two types of spanning trees:

- **Internal spanning tree (IST):** This is the spanning tree that runs in an MST region. Within each MST region, MST maintains multiple spanning-tree instances. Instance 0 is a special instance for a region known as the IST. All other MST instances are numbered from 1 to 4094.
- **Common and internal spanning tree (CIST):** This is a collection of the ISTs in each MST region and the common spanning tree (CST) that interconnects the MST regions and single spanning trees. The spanning tree computed in a region appears as a subtree in the CST that encompasses the entire switched domain. The CIST is formed by the spanning-tree algorithm running among switches that support the IEEE 802.1w, IEEE 802.1s, and IEEE 802.1D standards. The CIST inside an MST region is the same as the CST outside a region.

Example 1.22 shows the configuration and verification of MST. In this example, the MST region is named ExamCram with revision number 1. There are two instances. Instance 1 maps to VLANs 10 and 20, and instance 2 maps to VLANs 30 and 40. You change the mode to MST by using the command **spanning-tree mode mst**. You verify the instances and their mappings by using the command **show spanning-tree mst configuration**.

EXAMPLE 1.22 **Configuring and Verifying MST**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# spanning-tree mst configuration
SW1(config-mst)# name ExamCram
SW1(config-mst)# revision 1
SW1(config-mst)# instance 1 vlan 10,20
SW1(config-mst)# instance 2 vlan 30,40
SW1(config-mst)# show pending
Pending MST configuration
Name      [ExamCram]
Revision  1     Instances configured 3
```

```
Instance  Vlans mapped
--------  ---------------------------------------------------------------
0         1-9,11-19,21-29,31-39,41-4094
1         10,20
2         30,40
---------------------------------------------------------------------------
SW1(config-mst)# exit
SW1(config)# spanning-tree mode mst
SW1(config)# end
SW1#
SW1# show spanning-tree mst configuration
Name      [ExamCram]
Revision  1     Instances configured 3

Instance  Vlans mapped
--------  ---------------------------------------------------------------
0         1-9,11-19,21-29,31-39,41-4094
1         10,20
2         30,40
---------------------------------------------------------------------------
SW1#

SW3#
SW3# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW3(config)# spanning-tree mst configuration
SW3(config-mst)# name ExamCram
SW3(config-mst)# revision 1
SW3(config-mst)# instance 1 vlan 10,20
SW3(config-mst)# instance 2 vlan 30,40
SW3(config-mst)# show pending
Pending MST configuration
Name      [ExamCram]
Revision  1     Instances configured 3

Instance  Vlans mapped
--------  ---------------------------------------------------------------
0         1-9,11-19,21-29,31-39,41-4094
1         10,20
2         30,40
---------------------------------------------------------------------------
SW3(config-mst)# exit
SW3(config)# spanning-tree mode mst
SW3(config)# end
SW3#
```

```
SW3#
SW3# show spanning-tree mst configuration
Name      [ExamCram]
Revision  1     Instances configured 3

Instance  Vlans mapped
--------  ---------------------------------------------------------------
0         1-9,11-19,21-29,31-39,41-4094
1         10,20
2         30,40
---------------------------------------------------------------------------
SW3#
```

You can configure path cost and port priority with MST as follows:

- **spanning-tree mst** *instance-id* **port-priority** *priority***:** For *instance-id*, you can specify a single instance, a range of instances separated by a hyphen, or a series of instances separated by a comma; the range is 0 to 4094. For *priority*, the range is 0 to 240, in increments of 16, and the default is 128. The lower the number, the higher the priority. You can verify your entries with the **show spanning-tree mst** *instance-id* command.
- **spanning-tree mst** *instance-id* **cost** *cost***:** For *cost*, the range is 1 to 200,000,000; the default value is derived from the media speed of the interface. You can verify your entries with the **show spanning-tree mst** *instance-id* command.

You can set switch priority, as shown earlier in Example 1.16. However, in this case, you set the priority per MST instance and not per individual VLAN. You use the **spanning-tree mst** *instance-id* **root primary** and **spanning-tree mst** *instance-id* **root secondary** global configuration commands to modify the switch priority. Alternatively, if you need to configure a specific priority, you use the **spanning-tree mst** *instance-id* **priority** *priority* command. You can verify your configuration by using the **show spanning-tree mst** *instance-id* command.

Example 1.23 achieves load sharing by making different switches the root bridge for different MST instances. SW1 is configured to be the root bridge for instance 1 and the secondary root bridge for instance 2. SW3 is configured to be the root bridge for instance 2 and the secondary root bridge for instance 1. The command **show spanning-tree mst 1** shows this in the output.

EXAMPLE 1.23 **Configuring and Verifying Root Bridge Priority for MST**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# spanning-tree mst 1 root primary
SW1(config)# spanning-tree mst 2 root secondary
SW1(config)# end
SW1#

SW3#
SW3# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW3(config)# spanning-tree mst 2 root primary
SW3(config)# spanning-tree mst 1 root secondary
SW3(config)# end
SW3#

SW1#
SW1# show spanning-tree mst 1

##### MST1    vlans mapped:   10,20
Bridge        address 5254.001a.37c2  priority   24577 (24576 sysid 1)
Root          this switch for MST1

Interface        Role Sts Cost      Prio.Nbr Type
---------------- ---- --- --------- -------- --------------------------
Gi0/0            Desg FWD 20000     128.1    P2p

SW1# show spanning-tree mst 2

##### MST2    vlans mapped:   30,40
Bridge        address 5254.001a.37c2  priority   28674 (28672 sysid 2)
Root          address 5254.000b.9f89  priority   24578 (24576 sysid 2)
              port    Gi0/0           cost       20000     rem hops 19

Interface        Role Sts Cost      Prio.Nbr Type
---------------- ---- --- --------- -------- --------------------------
Gi0/0            Root FWD 20000     128.1    P2p
SW1#

SW3#
SW3# show spanning-tree mst 1

##### MST1    vlans mapped:   10,20
Bridge        address 5254.000b.9f89  priority   28673 (28672 sysid 1)
Root          address 5254.001a.37c2  priority   24577 (24576 sysid 1)
              port    Gi0/0           cost       20000     rem hops 19
```

```
Interface        Role Sts Cost      Prio.Nbr Type
---------------- ---- --- --------- -------- --------------------------
Gi0/0            Root FWD 20000     128.1    P2p

SW3# show spanning-tree mst 2

##### MST2    vlans mapped:   30,40
Bridge        address 5254.000b.9f89  priority   24578 (24576 sysid 2)
Root          this switch for MST2

Interface        Role Sts Cost      Prio.Nbr Type
---------------- ---- --- --------- -------- --------------------------
Gi0/0            Desg FWD 20000     128.1    P2p
SW3#
```

You saw earlier in the chapter how to configure the hello time, max age time, and forward delay time for each VLAN. The timers can be set globally for all MST instances:

- **spanning-tree mst hello-time** *seconds*: Verification is done with the **show spanning-tree mst** command.
- **spanning-tree mst forward-time** *seconds*: Verification is done with the **show spanning-tree mst** command.
- **spanning-tree mst max-age** *seconds*: Verification is done with the **show spanning-tree mst** command.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. On a non-root switch, which port is the closest to the root switch?
   - ❍ **A.** Root port
   - ❍ **B.** Designated port
   - ❍ **C.** Non-designated port
   - ❍ **D.** Forwarding port

2. How many bits are long method spanning-tree path cost values?
   - ❍ **A.** 8 bits
   - ❍ **B.** 16 bits
   - ❍ **C.** 32 bits
   - ❍ **D.** 64 bits

3. Which of the following is one of the benefits of using Multiple Spanning Protocol (MST)?
   - ❍ **A.** Increase in the number of spanning-tree instances
   - ❍ **B.** Reduction in the number of forwarding paths
   - ❍ **C.** Reduction in the number of spanning-tree instances
   - ❍ **D.** Reduction in spanning-tree convergence time

4. Which command is used to disable Spanning Tree Protocol PortFast on a particular interface after it is enabled globally?
   - ❍ **A. spanning-tree disable portfast**
   - ❍ **B. spanning-tree portfast disable**
   - ❍ **C. no spanning-tree portfast**
   - ❍ **D. no spanning-tree portfast enable**

# Answers

1. **A** is correct. The root port is the port that receives the best BPDU in terms of the path cost. It is the port that is the closest to the root switch in terms of path cost.
2. **C** is correct. The long spanning-tree path cost method uses 32-bit port cost values that are assigned to each port using a formula that is based on the port bandwidth.
3. **C** is correct. Multiple Spanning Tree Protocol (MST) enables you to map multiple VLANs to the same spanning-tree instance. This aids in reducing the number of spanning-tree instances needed to support a large number of VLANs.
4. **B** is correct. When PortFast is enabled globally, and you need to disable it on a specific port, you can use the command **spanning-tree portfast disable** to disable it on that specific port.

# EtherChannels

This section covers the configuration and troubleshooting of Layer 2 and Layer 3 EtherChannels.

Port channels, EtherChannels, and port aggregation all refer to the same group of technologies that enable the bonding of multiple physical links into a virtual link. Whereas port channels and aggregation of ports are general terms, EtherChannel is a Cisco brand name for its implementation of this technology. It provides automatic recovery for the loss of a link by redistributing the load across the remaining links. If a link fails, EtherChannel redirects traffic from the failed link to the channel's remaining links, without the network administrator's intervention.

Each EtherChannel can consist of up to eight compatibly configured ports. EtherChannel can be used to aggregate access, trunk, and routed ports. All ports in each EtherChannel must be configured as either Layer 2 or Layer 3 ports. The EtherChannel Layer 3 ports are made up of routed ports. Routed ports are physical ports configured to be in Layer 3 mode using the **no switchport** interface configuration command.

EtherChannel or port channel can be configured using one of these methods: using Port Aggregation Protocol (PAgP) or Link Aggregation Control Protocol (LACP) mode or using the ON mode. Both ends of the link need to be configured with the same method:

- **PAgP or LACP mode:** When you configure one end of an EtherChannel in either PAgP or LACP mode, the system negotiates with the other end of the channel to determine which ports should become active. If the remote port cannot negotiate an EtherChannel, the local port is put into an independent state and continues to carry data traffic, as would any other single link. The port configuration does not change, but the port fails to participate in the EtherChannel.
- **ON mode:** When you configure an EtherChannel in the ON mode, no negotiations take place. The switch forces all compatible ports to become active in the EtherChannel. The other end of the channel (on the other switch) must also be configured in the ON mode; otherwise, packet loss can occur.

Care should be taken when using the ON mode. This is a manual configuration, and ports on both ends of the EtherChannel must have the same configuration. If the group is misconfigured, packet loss or spanning-tree loops can occur.

EtherChannels provide a number of benefits:

- Optimized bandwidth usage
- Improved network convergence
- Spanning-tree mitigation
- Resilience against physical link failures

Let's now look at the two protocols that are used by EtherChannel to dynamically form port channel bundles: LACP and PAgP.

## LACP

Link Aggregation Control Protocol (LACP) is part of the IEEE 802.1AX specification. Because LACP is an IEEE standard, it can be used to facilitate port channels in mixed switch environments. LACP checks for configuration consistency and manages link additions and failures between two switches. It ensures that when you create a port channel, all the ports have the same type of configuration speed, duplex setting, and VLAN information.

**ExamAlert**

You should know the various combinations for forming an LACP EtherChannel for the ENCOR exam.

LACP has two modes:

- **Passive mode:** Passive mode places a port into a passive negotiating state, in which the port responds to LACP packets that it receives but does not start LACP packet negotiation. This mode minimizes the transmission of LACP packets. Passive mode is useful when you do not know whether the remote system or partner supports LACP.
- **Active mode:** Active mode places a port into an active negotiating state in which the port starts negotiations with other ports by sending LACP packets.

Both modes allow LACP to negotiate between ports to determine if they can form a port channel. This result is based on criteria such as the port speed and the trunking state.

Ports can form an EtherChannel when they are in different LACP modes as long as the modes are compatible at the other end of the link. For example:

- A port in **active** mode can form an EtherChannel with another port in **active** mode or **passive** mode.
- A port in **passive** mode cannot form an EtherChannel with another port in **passive** mode because neither port starts LACP negotiation.

Table 1.5 shows which combinations can and can't form an LACP EtherChannel.

TABLE 1.5 **Combinations That Can Form an LACP EtherChannel**

| | Passive | Active | On |
|---|---|---|---|
| **Passive** | Not formed | Formed | Not formed |
| **Active** | Formed | Formed | Not formed |
| **On** | Not formed | Not formed | Formed |

Because LACP is the protocol commonly used for forming EtherChannel bundles, it is used in the examples in this section.

You use the **channel-group** *interface* configuration command to dynamically create a port channel logical interface. The channel group number identifies the logical port channel. This number needs to be unique for each EtherChannel you create. Different Catalyst platforms support different numbers of EtherChannels. For example, the Catalyst 3850 supports a maximum of 128 EtherChannels.

Example 1.24 shows the use of the command **channel-group** *number* to bind the logical interface 1 to the physical interfaces GigabitEthernet 0/0 and 0/1. The **channel-group 1 mode** *mode* command specifics the LACP mode passive on SW1 and the LACP mode active on SW2. Shortly after configuration of SW2 physical ports in LACP active mode, you can see that the EtherChannel changes state to up.

EXAMPLE 1.24 **Configuring Layer 2 EtherChannel**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# interface range GigabitEthernet 0/0-1
```

```
SW1(config-if-range)# channel-group 1 mode passive
Creating a port-channel interface Port-channel 1

SW2#
SW2# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW2(config)# interface range GigabitEthernet 0/0-1
SW2(config-if-range)# channel-group 1 mode active
Creating a port-channel interface Port-channel 1

SW2(config-if-range)#
*May 10 17:03:46.277: %LINK-3-UPDOWN: Interface Port-channel1, changed
 state to up
*May 10 17:03:47.277: %LINEPROTO-5-UPDOWN: Line protocol on Interface
 Port-channel1, changed state to up
```

Example 1.25 shows the verification that the port channel has been established. The command **show etherchannel summary** shows the status of all configured EtherChannels, along with their dynamic aggregation protocol, if one was used. Note in this output that port channel 1 is in the SU state (which indicates Layer 2 and in use), LACP was used for negotiation, and the physical interfaces are in the P state (that is, bundled in port channel). Also, the command **show interface port-channel 1** indicates that the port channel is showing the consolidated bandwidth of the two 1 Gbps physical interfaces.

EXAMPLE 1.25 **Verifying Layer 2 EtherChannel**

```
SW2#
SW2# show etherchannel summary
Flags:  D - down        P - bundled in port-channel
        I - stand-alone s - suspended
        H - Hot-standby (LACP only)
        R - Layer3      S - Layer2
        U - in use      N - not in use, no aggregation
        f - failed to allocate aggregator

        M - not in use, minimum links not met
        m - not in use, port not aggregated due to minimum links not
            met
        u - unsuitable for bundling
        w - waiting to be aggregated
        d - default port

        A - formed by Auto LAG
```

```
Number of channel-groups in use: 1
Number of aggregators:           1

Group  Port-channel  Protocol    Ports
------+-------------+-----------+-----------------------------------
  -----------
1      Po1(SU)          LACP      Gi0/0(P)    Gi0/1(P)

SW2# show interface port-channel 1
Port-channel1 is up, line protocol is up (connected)
  Hardware is EtherChannel, address is 5254.0009.1714 (bia
 5254.0009.1714)
  MTU 1500 bytes, BW 2000000 Kbit/sec, DLY 10 usec,
  <... output omitted ...>
SW2#
```

Example 1.26 shows how to create a Layer 3 port channel. First, because the port channel will be a routed port, you need to use the command **no switchport** on the two physical interfaces that will be part of this port channel. You then use the command **channel-group 2** to bind the logical interface 2 to the physical interfaces GigabitEthernet 0/2 and 0/3. The **channel-group 2 mode** *active* command specifies that SW2 should initiate the formation of an EtherChannel using LACP.

EXAMPLE 1.26 **Configuring Layer 3 EtherChannel**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# interface range GigabitEthernet 0/2-3
SW1(config-if-range)# no switchport
SW1(config-if-range)# channel-group 2 mode passive
Creating a port-channel interface Port-channel 2
SW2#
SW2# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW2(config)# interface range GigabitEthernet 0/2-3
SW2(config-if-range)# no switchport
SW2(config-if-range)# channel-group 2 mode active
Creating a port-channel interface Port-channel 2
 SW2(config-if-range)#
*May 10 17:34:09.494: %LINK-3-UPDOWN: Interface Port-channel2, changed
 state to up
*May 10 17:34:10.496: %LINEPROTO-5-UPDOWN: Line protocol on Interface
 Port-channel2, changed state to up
SW2(config-if-range)# end
SW2#
```

Example 1.27 shows the verification that the port channel has been established. The **show etherchannel summary** output shows that port channel 2 is in RU state (that is, Layer 3 and in use), LACP was used for negotiation, and the physical interfaces are in the P state (that is, bundled in port channel).

EXAMPLE 1.27 **Verifying Layer 3 EtherChannel**

```
SW2# show etherchannel summary
Flags:  D - down        P - bundled in port-channel
        I - stand-alone s - suspended
        H - Hot-standby (LACP only)
        R - Layer3      S - Layer2
        U - in use      N - not in use, no aggregation
        f - failed to allocate aggregator

        M - not in use, minimum links not met
        m - not in use, port not aggregated due to minimum links not
            met
        u - unsuitable for bundling
        w - waiting to be aggregated
        d - default port

        A - formed by Auto LAG


Number of channel-groups in use: 2
Number of aggregators:           2

Group  Port-channel  Protocol    Ports
------+-------------+-----------+-----------------------------------
1      Po1(SU)         LACP      Gi0/0(P)    Gi0/1(P)
2      Po2(RU)         LACP      Gi0/2(P)    Gi0/3(P)

SW2#
```

# PAgP

Port Aggregation Protocol (PAgP) is a Cisco-proprietary protocol that can run only on Cisco switches. PAgP facilitates the automatic creation of EtherChannels by exchanging PAgP packets between Ethernet ports. Much like LACP, PAgP checks for configuration consistency. PAgP groups ports with the same speed, duplex mode, native VLAN, VLAN range, and trunking status and type.

**ExamAlert**

The ENCOR exam is likely to cover the various combinations for forming a PAgP EtherChannel.

PAgP has two modes:

- **Auto mode:** In auto mode, a port is placed into a passive negotiating state, in which the port responds to PAgP packets it receives but does not start PAgP packet negotiations. This setting minimizes the transmission of PAgP packets.
- **Desirable mode:** In desirable mode, a port is placed into an active negotiating state, in which the port starts negotiations with other ports by sending PAgP packets.

Ports can form an EtherChannel when they are in different PAgP modes as long as the modes are compatible at the other end. For example:

- A port in **desirable** mode can form an EtherChannel with another port in **desirable** or **auto** mode.
- A port in **auto** mode can form an EtherChannel with another port in **desirable** mode.

A port in **auto** mode cannot form an EtherChannel with another port in **auto** mode because neither port starts PAgP negotiations.

Table 1.6 shows the combinations that can and cannot form a PAgP EtherChannel.

TABLE 1.6 **Combinations That Can Form a PAgP EtherChannel**

| | Auto | Desirable | On |
|---|---|---|---|
| **Auto** | Not formed | Formed | Not formed |
| **Desirable** | Formed | Formed | Not formed |
| **On** | Not formed | Not formed | Formed |

EtherChannel supports several options for the load balancing of traffic across links in a bundle. However, traffic is not necessarily distributed equally across the links in a bundle. These are the EtherChannel load balancing methods:

- **dst-ip:** Destination IP address
- **dst-mac:** Destination MAC address

- **src-dst-ip:** Source and destination IP address
- **src-dst-mac:** Source and destination MAC address
- **src-ip:** Source IP address
- **src-mac:** Source MAC address
- **src-port:** Source port number
- **dst-port:** Destination port number
- **src-dst-port:** Source and destination port number

Changing the load balancing method applies to all EtherChannels configured on a switch, as shown in Example 1.28. You configure the load-balancing and forwarding method by using the **port-channel load-balance** global configuration command. You use the **show etherchannel load-balance** command to verify how a switch will load balance traffic.

EXAMPLE 1.28 **Configuring and Verifying EtherChannel Load Balancing**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# port-channel load-balance ?
  dst-ip        Dst IP Addr
  dst-mac       Dst Mac Addr
  src-dst-ip    Src XOR Dst IP Addr
  src-dst-mac   Src XOR Dst Mac Addr
  src-ip        Src IP Addr
  src-mac       Src Mac Addr

SW1(config)# end
SW1#
SW1# show etherchannel load-balance
EtherChannel Load-Balancing Configuration:
        src-dst-ip

EtherChannel Load-Balancing Addresses Used Per-Protocol:
Non-IP: Source XOR Destination MAC address
  IPv4: Source XOR Destination IP address
  IPv6: Source XOR Destination IP address

SW1#
```

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which of the following statements is correct in relationship to implementing a port channel between two switches with LACP as the negotiation protocol?
   - ❍ **A.** A port in active mode can form a port channel with another port that is in passive mode.
   - ❍ **B.** A port in passive mode can form a port channel with another port that is in passive mode.
   - ❍ **C.** A port in ON mode can form a port channel with another port that is in active mode.
   - ❍ **D.** A port in ON mode can form a port channel with another port that is in passive mode.

2. EtherChannel allows for the aggregation of which of the following types of ports? (Choose all that apply.)
   - ❍ **A.** Loopback
   - ❍ **B.** Routed
   - ❍ **C.** Access
   - ❍ **D.** Trunk

3. What is the maximum number of links that can be bundled with EtherChannel?
   - ❍ **A.** 2
   - ❍ **B.** 4
   - ❍ **C.** 8
   - ❍ **D.** 16

4. Which of the following is not a benefit of EtherChannel?
   - ❍ **A.** Optimized bandwidth usage
   - ❍ **B.** Improved network convergence
   - ❍ **C.** Resilience against physical link failures
   - ❍ **D.** PAgP protocol support for forming bundles with non-Cisco switches

# Answers

1. **A** is correct. A port in **active** mode can form an EtherChannel with another port in **active** mode or **passive** mode.
2. **B**, **C**, and **D** are correct. EtherChannel can be used to bundle access, trunk, and routed ports.
3. **C** is correct. Each EtherChannel can consist of up to eight compatibly configured ports.
4. **D** is correct. Port Aggregation Protocol (PAgP) is a Cisco-proprietary protocol that can be run only on Cisco switches.

# Review Questions

1. Which version of VLAN Trunking Protocol (VTP) can be used to propagate Multiple Spanning Tree Protocol (MST) database information?
   - ❍ **A.** VTP Version 1
   - ❍ **B.** VTP Version 2
   - ❍ **C.** VTP Version 3
   - ❍ **D.** None of the above, as VTP cannot be used for propagating MST information

2. In which DTP mode does the interface try to actively negotiate the formation of a trunk?
   - ❍ **A.** Dynamic desirable
   - ❍ **B.** Dynamic auto
   - ❍ **C.** Trunk
   - ❍ **D.** On

3. Which Spanning Tree Protocol feature causes a port to bypass the listening and learning states and enter the forwarding state immediately?
   - ❍ **A.** Root Guard
   - ❍ **B.** PortFast
   - ❍ **C.** Loop Guard
   - ❍ **D.** Bridge Assurance

4. A Cisco switch Spanning Tree Protocol priority can be configured in increments of __?
   - ❍ **A.** 10
   - ❍ **B.** 1024
   - ❍ **C.** 2048
   - ❍ **D.** 4096

5. True or false: Port Aggregation Protocol (PAgP) is a link aggregation protocol that is Cisco proprietary.
   - ❍ **A.** True
   - ❍ **B.** False

6. True or false: Layer 3 EtherChannels requires the **no switchport** interface configuration command to be run on all physical interfaces that will be part of the EtherChannel bundle.
   - ❍ **A.** True
   - ❍ **B.** False

# Answers to Review Questions

1. **C** is correct. VTP Version 3 can be used to propagate MST database information.
2. **A** is correct. In dynamic auto mode, the interface actively tries to convert the link to a trunk link. It becomes a trunk interface if the other end is set to trunk, desirable, or auto mode.
3. **B** is correct. A port with PortFast enabled bypasses the listening and learning states and enters forwarding state immediately.
4. **D** is correct. Switch priority can be configured with a value between 0 to 61440, in increments of 4096.
5. **A** is correct. Port Aggregation Protocol (PAgP) is the Cisco-proprietary dynamic link aggregation protocol.
6. **A** is correct. Because a Layer 3 EtherChannel is a routed port, you need to use the command **no switchport** on all of the physical interfaces that will be part of the EtherChannel bundle.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*
- *CCNP and CCIE Enterprise Core & CCNP Advanced Routing Portable Command Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers Layer 3 interior gateway protocols (IGPs).

CHAPTER 2

# Understanding Layer 3: IGPs

**This chapter covers the following official ENCOR 350-401 exam objectives:**

- 3.2 Layer 3
- 3.2.a Compare routing concepts of EIGRP and OSPF (advanced distance vector vs. linked state, load balancing, path selection, path operations, metrics)
- 3.2.b Configure and verify simple OSPF environments, including multiple normal areas, summarization, and filtering (neighbor adjacency, point-to-point and broadcast network types, and passive interface)

This chapter covers IP routing essentials as well as routing concepts associated with the Enhanced Interior Gateway Routing Protocol (EIGRP) and Open Shortest Path First (OSPF) interior gateway protocols (IGPs). It provides an overview of routing protocols, routing path selection, static routing, and other routing concepts. The second section looks at the fundamentals of the EIGRP routing protocol, its path metric calculation, failure detection and timers, and route summarization. The final section of this chapter reviews the fundamentals of the OSPF routing protocol along with its configuration and verification, default route advertisement, and OSPF optimization. This section also examines advanced OPSF concepts, including multiple areas and summarization of routes. This chapter concludes by looking at the fundamentals of OSPFv3 and IPv4 support in OSPFv3.

**This chapter covers the following technology topics:**

- IP Routing Essentials
- Enhanced Interior Gateway Routing Protocol (EIGRP)
- Open Shortest Path First (OSPF)

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. What is the purpose of a floating static route?
2. What is the default AD (administrative distance) that OSPF uses?
3. What does the EIGRP topology table contain?
4. What is the hello timer interval that EIGRP uses on a slow interface?
5. What is the main purpose of a virtual link in an OSPF implementation?
6. In OSPFv3, the AllDRouters designated router (DR) is represented by what address?

### Answers

1. A floating static route is a static route that a router uses as a backup to a primary or dynamic route.
2. 110
3. The EIGRP topology table contains all of the prefixes that are advertised in that particular EIGRP autonomous system.
4. 60 seconds
5. A virtual link in OSPF provides a disconnected area with a logical path to the backbone (Area 0).
6. FF02::06

# IP Routing Essentials

This section provides an overview of IP routing protocols and differences between distance vector and link-state protocols. It also reviews the path selection logic that routers use to select the best route to install in a routing table. Finally, this section looks at concepts related to static routing.

The primary function of a router is to move packets from one network to another. Networks that are not directly connected to a router are learned through the configuration of static routes or the configuration of dynamic IP routing protocols. A number of routing protocols are found in most routing platforms:

- Routing Information Protocol Version 2 (RIPv2)
- Enhanced Interior Gateway Routing Protocol (EIGRP)

- Open Shortest Path First (OSPF)
- Intermediate System-to-Intermediate System (IS-IS)
- Border Gateway Protocol (BGP)

The first four protocols in this list are designed to route within an autonomous system (AS). An AS is a network of routers and related systems that are managed under a common administrative boundary. These first four routing protocols are considered interior gateway protocols (IGPs), and BGP is considered an exterior gateway protocol (EGP) that is used for routing between different autonomous systems. This is known as an exterior BGP (eBGP) session. BGP can also be used for routing within an autonomous system, and this is known as an interior BGP (iBGP) session. BGP is reviewed in Chapter 3, "Understanding Layer 3: BGP."

# Routing Algorithms

Routing protocols can be classified as using several types of algorithms:

- **Distance vector algorithms (also known as Bellman–Ford algorithms):** Each router sends all or some portion of its routing table only to directly connected devices. In essence, distance vector algorithms send larger routing updates only to directly connected devices, and these algorithms know only about their directly connect devices. Distance vector protocols have the advantage of using fewer CPU and memory resources but do not scale well and therefore are useless in most modern networks that require scalability. RIP is a distance vector protocol.
- **Advanced distance vector algorithms:** The diffusing update algorithm (DUAL) is an advanced distance vector algorithm. Some of the advantages that an advanced distance vector algorithm provides over distance vector algorithms include rapid convergence, use of hello mechanisms, formation of neighborships, and the ability to balance traffic over equal- and unequal-metric paths. EIGRP is an advanced distance vector protocol.
- **Link-state algorithms (also known as shortest path first algorithms):** With a link-state algorithm, each router builds a picture of the network and calculates the shortest path to each known destination. The best path is then calculated using the shortest path first (SPF) algorithm, also called Dijkstra's algorithm, and the best path to every destination is then installed into the routing table. OSPF and IS-IS are two link-state routing protocols.

- **Path vector algorithms:** The path vector protocol is similar to distance vector protocols, but instead of looking at the distance to figure out the best loop-free path, it looks at various BGP path attributes. These BGP path attributes include autonomous system path (AS_Path), multi-exit discriminator (MED), origin, next hop, local preference, atomic aggregate, and aggregator. A path vector protocol can provide loop-free paths by keeping track of each autonomous system that the routing advertisement traverses.

# Path Selection

For path selection, a router identifies the path that a packet should take by looking at the prefix length that it sees in the forwarding information base (FIB). The FIB is programmed through the routing table or routing information base (RIB). The routes in the RIB are gathered from the routing protocol processes. The three main components of path selection are as follows:

- **Prefix length:** Prefix length is the number of leading binary bits in the subnet mask that are in the *on* position.
- **Administrative distance (AD):** Administrative distance is the rating of the trustworthiness of a routing information source. Table 2.1 shows the default administrative distance values for routes learned from different sources.

> **ExamAlert**
>
> For the ENCOR exam, it would be helpful to memorize the different routing sources and the default administrative distance for each of them.

TABLE 2.1 **Default Administrative Distance Values**

| Route Source | Default Administrative Distance |
|---|---|
| Connected interface | 0 |
| Static route | 1 |
| EIGRP summary route | 5 |
| BGP | 20 |
| Internal EIGRP | 90 |
| OSPF | 110 |

| Route Source | Default Administrative Distance |
|---|---|
| IS-IS | 115 |
| RIP | 120 |
| EIGRP external route | 170 |
| Internal and local BGP | 200 |
| Unknown | 255 |

- **Metrics:** A metric is a standard of measurement (for example, path bandwidth) used by routing algorithms to determine the optimal path to a destination.

Path selection is an important part of understanding how a router forwards packets, and we take a closer look at prefix length, AD, and metrics next.

ExamAlert

Path selection is an important topic for the ENCOR exam, and you should have a complete understanding of it.

## Prefix Length

If three routes have different network prefix lengths (network masks), the three routes are considered unique and are entered into the routing table. For example, say that a router has the following three routes with various prefix lengths to reach a destination:

- 10.0.5.0/24
- 10.0.5.0/26
- 10.0.5.0/28

The packet forwarding logic determines which one of the three to use. The longest prefix match always wins and is the route that is installed in the routing table. In this case, that will be 10.0.5.0/28. If the packet needs to be forwarded, the route is chosen based on the prefix length. For example, if the packet needs to be forwarded to 10.0.5.14, the router matches on three routes as the packet fits into all three IP address ranges. The packet will be forwarded to the next hop on the outgoing interface that is associated with the longest matched prefix (10.0.5.0/28). On the other hand, if the packet needs to be forwarded to

10.0.5.42, the router matches on both the 10.0.5.0/24 and the 10.0.5.0/26 prefixes. The packet will be forwarded to the next hop on the outgoing interface that is associated with the longest matched prefix (10.0.5.0/26 in this case).

## Administrative Distance

An AD is an integer from 0 to 255, and the higher the value, the lower the trust rating. An administrative distance of 255 means the routing information source cannot be trusted and should be ignored. The logic related to how a router accepts or reject routes in the RIB is as follows:

- If the route does not exist in the RIB, it is accepted.
- If the route exists in the RIB, the AD is compared. If the AD of the route in the RIB is lower than the AD of the second route, then the route is rejected.
- If the route exists in the RIB, the AD is compared. If the AD of the route in the RIB is higher than the AD of the second route, then the route is accepted, and the current source protocol is notified of the removal of the existing entry in the RIB.

Changing the AD on a routing protocol is highly discouraged as it can have severe consequences, such as routing loops and other odd behaviors in the network.

## Metrics

Metrics vary from one routing protocol to another. IGPs typically prefer internally learned routes over external routes; in addition, IGPs select the path with the lowest metric. If the routing protocol identifies multiple paths as the best path (that is, with the same metrics), the router installs the maximum number of paths allowed per destination. This is considered equal-cost multipathing, and it provides load sharing across all links.

EIGRP can be configured to support unequal-cost load balancing, where it installs multiple routes with different path metrics. This is not enabled by default. Unequal-cost load balancing allows for traffic to be transmitted out of a router's interface in a ratio relative to the path metric of the interface.

**ExamAlert**

Understanding the process on a router is critical for both the ENCOR exam and real-world deployments.

# Static Routing

A network administrator manually establishes static routing to provide precise control over routing, and static routing does not change unless the network administrator alters it. Static routes are simple to create and work well in environments where network traffic is relatively predictable and where the network design is fairly simple. Static routing may create an administrative burden as the number of routers and network segments grows. Because static routing systems cannot react to network changes, they generally are considered unsuitable for large, constantly changing networks. You typically use a dynamic routing protocol that adjusts to changing network circumstances by analyzing incoming routing update messages as you scale the network.

Because routers do not communicate states when using static routes, there is no network intelligence. For example, as a link goes down, other routers would not be made aware of the change or that the network path is no longer accessible. However, there are some cases where static routes may be useful:

- When a router has limited CPU and memory resources and cannot use a dynamic routing protocol
- When you need to supersede the routes learned and chosen from a dynamic routing protocol

Static routes can be classified into the following types:

- **Directly attached static routes:** Static routes can reference the outbound interface of a router. A static route that uses only the outbound next-hop interface is a directly attached static route. The requirement for this to work is that the outbound interface needs to be in an *up* state before the route can be installed into the RIB. You can configure directly attached static routes by using the command **ip route** *network subnet-mask next-hop-interface-id*.
- **Recursive static routes:** A recursive static route is a route whose next hop and the destination network are covered by another learned route in the RIB. The recursive static route feature allows you to install recursive static routes in the RIB. It requires the route's next-hop address to be reachable via some other existing route in the routing table so that the static route can be installed in the RIB. A recursive static route specifies the IP address of the next-hop address. You configure recursive static routes by using the command **ip route** *network subnet-mask next-hop-ip*.
- **Fully specified static routes:** A fully specified static route is a route that specifies both an interface and a next-hop IP address. You can use a fully

specified static route when the output interface is a multi-access interface, and you need to identify the next-hop address. The next-hop address must be directly attached to the specified output interface. You configure fully specified static routes by using the command **ip route** *network subnet-mask interface-id next-hop-ip*.

- **Floating static routes:** A floating static route is a static route that the router uses as a backup to a primary or dynamic route. You must configure a floating static route with a higher administrative distance than the primary or dynamic route that it is backing up. This causes the router to prefer a primary or dynamic route to a floating static route. You can then use a floating static route as a replacement if the primary or dynamic route is lost. If no administrative distance is configured, the default AD (that is, 1) is used. To configure a floating static route, you create a static route and specify a higher AD at the end, as in these examples:
  - **ip route** *network subnet-mask next-hop-interface-id* **205**
  - **ip route** *network subnet-mask next-hop-ip* **210**

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which type of static route requires that the next hop and the destination network be covered by another learned route in the RIB?
   - ❍ **A.** Fully specified static route
   - ❍ **B.** Directly attached static route
   - ❍ **C.** Floating static route
   - ❍ **D.** Recursive static route

2. Which of the following routing protocols use a link-state algorithm to calculate the best path? (Choose two.)
   - ❍ **A.** RIP
   - ❍ **B.** EIGRP
   - ❍ **C.** OSPF
   - ❍ **D.** IS-IS
   - ❍ **E.** BGP

# Answers

1. **D** is correct. A recursive static route is a route whose next hop and the destination network are covered by another learned route in the RIB.
2. **C** and **D** are correct. OSPF and IS-IS use a link-state algorithm to calculate the best path.

# Enhanced Interior Gateway Routing Protocol (EIGRP)

This section reviews the first of three dynamic routing protocols. EIGRP is an enhanced distance vector routing protocol that relies on the diffusing update algorithm (DUAL) to calculate the shortest path to a network. EIGRP was initially a Cisco-proprietary protocol but was released to the Internet Engineering Task Force (IETF) through RFC 7868, ratified in May 2016.

EIGRP is ideal for large enterprise networks and is known for its ease of deployment and fast convergence capabilities. EIGRP maintains the advantages of traditional distance vector protocols while avoiding their disadvantages.

**ExamAlert**

It would be helpful to remember the advantages of EIGRP for the ENCOR exam.

Some of the advantages of using EIGRP are as follows:

- Eases the transition to IPv6 because it provides multi-address family support for both IPv4 and IPv6 networks.
- Provides superior scaling for an IGP for large Dynamic Multipoint Virtual Private Network (DMVPN) deployments.
- Provides extremely quick convergence times for changes in the network topology.
- Propagates only routing table changes, not the entire routing table, when a change occurs.
- Facilitates more efficient use of links through equal-cost multipathing (ECMP) and unequal-cost load sharing.

The following terms are associated with EIGRP:

- **Successor route:** The successor route is the route with the lowest-metric path to reach the destination network.
- **Successor:** A successor is the first next-hop router for the successor route.

- **Feasible distance (FD):** FD is the metric value for the lowest-metric path to reach a destination network. It is calculated using this formula:

$$\text{Metric} = \left[ (K1 \cdot BW + \frac{K2 \cdot BW}{256 - Load} + K3 \cdot Delay) \cdot \frac{K5}{K4 + Reliability} \right]$$

- **Reported distance (RD):** RD is the distance reported by a router to reach a particular network prefix. The RD value is the feasible distance for the advertising router.
- **Feasibility condition:** The feasibility condition states that for a route to be considered a backup route, the RD received for that route must be less than the FD that is calculated locally. This logic guarantees a loop-free path.
- **Feasible successor:** A feasible successor is not the current best route but a backup route that can be used if the successor route is lost. In other words, it is a route that satisfies the feasibility condition and is kept as a backup route. The feasibility condition guarantees that the backup route is loop free.

Let us now look at a simple example of enabling EIGRP on an interface and verifying the configuration.

Example 2.1 shows the process of enabling EIGRP between two routers. The process is started using the **router eigrp** *process-number* command. Process number 1 is used in this case. The **network** command is used to enable EIGRP on an interface in the 172.16.0.0/24 network, using the wildcard mask 0.0.0.255.

EXAMPLE 2.1 **Configuring EIGRP**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# router eigrp 1
R1(config-router)# network 172.16.0.0 0.0.0.255
R1(config-router)# end
R1#

R2#
R2# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R2(config)# router eigrp 1
```

```
R2(config-router)# network 172.16.0.0 0.0.0.255
R2(config-router)# end
R2#
```

Example 2.2 shows the verification that EIGRP has been enabled on the interface. The command **show ip eigrp interfaces** is used to verify this configuration.

EXAMPLE 2.2 **Verifying EIGRP**

```
R1#
R1# show ip eigrp interfaces
EIGRP-IPv4 Interfaces for AS(1)
                   Xmit Queue   PeerQ        Mean   Pacing Time   Multicast    Pending
Interface   Peers  Un/Reliable  Un/Reliable  SRTT   Un/Reliable   Flow Timer   Routes
Gi0/0         1         0/0        0/0          8        0/0           50           0
R1#
```

Let us now take a look at the various tables that are used by EIGRP: the neighbor table, the topology table, and the routing table.

# Neighbor Table

EIGRP does not rely on the periodic advertisement of all network prefixes, which is done with other routing protocols, such as RIP and OSPF. Neighbor relationship is achieved with low overhead by routers when they periodically send small hello packets. As long as hello packets are received, the Cisco software can determine whether a neighbor is alive and functioning. After the status of the neighbor is determined, neighboring devices can exchange routing information. Reliable Transport Protocol (RTP) is responsible for the guaranteed, ordered delivery of EIGRP packets to all neighbors. Some EIGRP packets (such as updates) must be sent reliably; this means that the packets require acknowledgment from the destination. Confirming that packets are received makes the transport method reliable. Update, query, and reply packets are deemed reliable, but hello and ACK packets do not require acknowledgment, and they could be unreliable.

The main reason EIGRP sends hello packets is to ensure that its neighbors are healthy and available. EIGRP hello packets are sent at intervals that are determined by the hello timer. The default hello timer is normally 5 seconds, but it is 60 seconds for slower interfaces. EIGRP also uses hold time, which is

the amount of time EIGRP deems the router as still reachable and functioning. The default hold time is three times the hello time interval. So, the default is 15 seconds, and it is 180 seconds on slower links. If the hold time reaches 0, the EIGRP process declares the neighbor unreachable and notifies DUAL of a topology change.

EIGRP uses five packet types, shown in Table 2.2, to communicate with other routers. It uses IP protocol number 88, uses multicast packets where possible, and uses unicast when necessary. Communication happens using the multicast group address 224.0.0.10.

TABLE 2.2 **EIGRP Packet Types**

| Packet Type | Packet Name | Function |
|---|---|---|
| 1 | Hello | Used for discovering EIGRP neighbors and for detecting when a neighbor is lost |
| 2 | Request | Used to get specific information from a neighbor |
| 3 | Update | Used to transmit routing and reachability information with EIGRP neighbors |
| 4 | Query | Used to search for other paths during convergence |
| 5 | Reply | Sent in response to a query packet |

Before routes can be processed and added to the RIB, EIGRP requires a neighbor relationship to be formed. Once a router hears an EIGRP hello packet, the router attempts to become a neighbor with the other router. These parameters must match for EIGRP neighborship to be established:

- Metric formula K values
- Primary subnet matches
- Autonomous system number (ASN) matches
- Authentication parameters

Example 2.3 shows the verification of the EIGRP neighbor table using the **show ip eigrp neighbors** command. The output shows the EIGRP neighbor table of R1. As you can see, there is one neighbor (172.16.0.2) here, and it happens to be R2 on interface Gi0/0.

EXAMPLE 2.3 **Verifying the EIGRP Neighbor Table with show ip eigrp neighbors**

```
R1#
R1# show ip eigrp neighbors
EIGRP-IPv4 Neighbors for AS(1)
H   Address          Interface       Hold Uptime   SRTT   RTO  Q  Seq
                                     (sec)         (ms)       Cnt Num
0   172.16.0.2       Gi0/0             12 00:15:43    4   100  0  4
R1#
```

Now that we have reviewed how EIGRP stores its neighbor information in a neighbor table, let us look at the EIGRP topology table.

# Topology Table

EIGRP contains a topology table that is an important component to DUAL and stores information to identify loop-free backup routes. It contains all of the prefixes that are advertised in a particular EIGRP autonomous system.

When an EIGRP router discovers a new neighbor, it sends updates about the routes that it knows and receives routes from that new neighbor. These are the updates that make up the topology table, which basically contains all of the destinations advertised by the neighboring routers; in other words, each router stores the routing table of its neighbors in the topology table.

The entries in the topology tables contain the following information:

- Network prefix
- EIGRP neighbors that advertise the prefix
- Metrics for each neighbor (including RD and hop count)
- Values used to calculate the metric (load, reliability, minimum bandwidth, and so on)

You can use the command **show ip eigrp topology [all-links]** to display the topology table. By default, the topology table only displays the successor and feasible successor routes, and the **all-links** keyword displays the paths that did not meet the feasible condition.

The output in Example 2.4 shows three routes and their successors. They are all in the passive (P) state. A route is considered *passive* when the router is not currently performing any recomputation on the route. When a route is *active*, it is undergoing recomputation or looking for a new successor.

Example 2.4 shows the output of the **show ip eigrp topology [all-links]** command.

EXAMPLE 2.4 **The show ip eigrp topology [all-links] Command**

```
R1# show ip eigrp topology all-links
EIGRP-IPv4 Topology Table for AS(1)/ID(10.0.0.1)
Codes: P - Passive, A - Active, U - Update, Q - Query, R - Reply,
       r - reply Status, s - sia Status

P 192.168.0.0/16, 1 successors, FD is 130816, serno 3
        via 172.16.0.2 (130816/128256), GigabitEthernet0/0
P 172.16.0.0/24, 1 successors, FD is 2816, serno 1
        via Connected, GigabitEthernet0/0
P 10.0.0.0/24, 1 successors, FD is 128256, serno 2
        via Connected, Loopback0
        via 172.16.0.2 (130816/128256), GigabitEthernet0/0
```

EIGRP does not use a single attribute in the metric for its routes, as other routing protocols (like RIP and OSPF) do. EIGRP uses a combination of five elements that relate to the physical characteristics of an interface to determine the metric: bandwidth, load, delay, reliability, and MTU.

**ExamAlert**

Knowing the EIGRP K values, components, and functions is critical for the ENCOR exam.

Table 2.3 shows the EIGRP metric components.

TABLE 2.3 **EIGRP Metric Components**

| K Value | Component | Description |
|---|---|---|
| K1 | Bandwidth | The minimum bandwidth of the route, in Kbps. It can be 0 or any positive integer. The bandwidth for the metric formula is scaled and inverted by using this formula: Scaled Bandwidth = 107/Minimum Bandwidth, in Kbps. |
| K2 | Load | The effective load of the route, expressed as a number from 0 to 255, where 255 is 100% loading. |
| K3 | Delay | The route delay, in tens of microseconds: Scaled delay = Delay/10. |

| K Value | Component | Description |
|---|---|---|
| K4 | Reliability | The likelihood of successful packet transmission, expressed as a number between 0 and 255, where 255 means 100% reliability and 0 means no reliability. |
| K5 | MTU | The minimum maximum transmission unit (MTU) size of the route, in bytes. It can be 0 or any positive integer. |

EIGRP Wide Metrics is used to facilitate interfaces with bandwidths above 1 Gbps and up to 4.2 Tbps. To allow EIGRP to perform path selection, the EIGRP composite cost metric formula was modified. The paths are selected based on the computed time, and the time that information takes to travel through links is measured in picoseconds. The formula is as follows:

$$\text{Metric} = [(K1 \times \text{Minimum Throughput} + \{K2 \times \text{Minimum Throughput}\} / 256 - \text{Load}) + (K3 \times \text{Total Latency}) + (K6 \times \text{Extended Attributes})] * [K5/(K4 + \text{Reliability})]$$

EIGRP Wide Metrics introduced K6 as an additional K value. K6 allows for extended attributes. There are currently two extended attributes: jitter and energy. With the calculation of larger bandwidths, EIGRP can no longer fit the computed metric into the 4-byte unsigned long value needed by the RIB. You set the RIB scaling factor for EIGRP by using the **metric rib-scale** command. When you configure the **metric rib-scale** command, all EIGRP routes in the RIB are cleared and replaced with the new metric values.

Example 2.5 shows the verification of the EIGRP metric weight settings using the **show ip protocol** command.

EXAMPLE 2.5 **Verifying an EIGRP Protocol**

```
R1#
R1# show ip protocol
*** IP Routing is NSF aware ***

Routing Protocol is "application"
  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
  Maximum path: 32
  Routing for Networks:
  Routing Information Sources:
```

```
    Gateway         Distance      Last Update
  Distance: (default is 4)

Routing Protocol is "eigrp 1"
  Outgoing update filter list for all interfaces is not set
  Incoming update filter list for all interfaces is not set
  Default networks flagged in outgoing updates
  Default networks accepted from incoming updates
  EIGRP-IPv4 Protocol for AS(1)
    Metric weight K1=1, K2=0, K3=1, K4=0, K5=0
    Soft SIA disabled
    NSF-aware route hold timer is 240
<... output omitted ...>
```

Now that we have examined the topology table and metrics, let's now look at what routes make it into the routing table.

# Routing Tables

A routing table holds the best route to each destination, and it is used to forward packets. From the topology table, the successor route is offered to the routing table. Normal routing principles apply to the routing table. For example, if a router running EIGRP learns of multiple routes to the same destination from different sources, it uses the administrative distance to determine which route to keep.

As you saw earlier in the chapter, EIGRP uses DUAL to achieve rapid convergence. If the primary route inside a routing table fails, then the best backup route is immediately added to the routing table. If that route is not appropriate or if the backup route is not found in the routing table, EIGRP queries its neighbor to discover an alternate route.

Example 2.6 shows the contents of the routing table on R1. The D indicates that it is an EIGRP route that is learned via the neighbor 172.16.0.2. It is using the default EIGRP AD of 90, and the metric is 130816.

EXAMPLE 2.6 **A Routing Table Showing an EIGRP Route**

```
R1# show ip route eigrp
Codes: L - local, C - connected, S - static, 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
       i - IS-IS, su - IS-IS summary, 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, H - NHRP,
l - LISP
       a - application route
       + - replicated route, % - next hop override, p - overrides from
PfR

Gateway of last resort is not set

D     192.168.0.0/16 [90/130816] via 172.16.0.2, 00:08:30,
GigabitEthernet0/0
R1#
```

Next, let us look at authentication between EIGRP neighbors.

# EIGRP Authentication

EIGRP can authenticate the packets sent between neighbors to ensure that a router accepts packets only from routers with the same pre-shared key. EIGRP provides MD5 authentication for routing updates. The MD5 keyed digest in each EIGRP packet prevents the introduction of unauthorized routing messages from unapproved sources. You can configure multiple keys with specific lifetimes. After you define the keychain with the key ID, key string, and lifetime parameters, you can set up authentication under a particular interface by using the following commands:

```
ip authentication mode eigrp autonomous-system md5
ip authentication key-chain eigrp autonomous-system key-chain
```

# EIGRP Named Mode

When you use the **router eigrp** command with the *virtual-instance-name* argument, you create an EIGRP configuration called the *EIGRP named configuration* or *EIGRP named mode*. An EIGRP named configuration does not create an EIGRP routing instance by itself. It is a base configuration that is required to define address-family configurations that are used for routing.

The way EIGRP has been configured so far in this chapter is considered classic mode. This is how EIGRP was configured for years. In classic mode, the EIGRP configurations are scattered across the router and interface modes.

Named mode allows for configurations to be entered in a hierarchical fashion under the router mode. Named mode allows for multiple address families and AS number combinations. It allows you to have similar configurations for IPv4 and IPv6.

Different routers configured for EIGRP in the traditional manner and configured using named mode are compatible; in other words, an EIGRP-speaking router configured in classic mode can form a neighborship with an EIGRP-speaking router configured in named mode.

Example 2.7 shows EIGRP named mode configuration.

EXAMPLE 2.7 **Configuring EIGRP Named Mode**

```
R1# config terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# router eigrp examcram
R1(config-router)# address-family ipv4 autonomous-system 100
R1(config-router-af)# network 172.16.1.0
R1(config-router-af)# metric weights 0 1 0 1 0 0 0
R1(config-router-af)# exit
```

Example 2.8 shows EIGRP named mode verification. As you can see, the EIGRP Wide Metrics feature only works in the EIGRP named mode configuration. (Note the presence of the K6 value.)

EXAMPLE 2.8 **Verifying EIGRP Named Mode**

```
R1# show ip protocol
*** IP Routing is NSF aware ***
Routing Protocol is "application"
  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
  Maximum path: 32
  Routing for Networks:
  Routing Information Sources:
    Gateway         Distance      Last Update
  Distance: (default is 4)

Routing Protocol is "eigrp 100"
  Outgoing update filter list for all interfaces is not set
  Incoming update filter list for all interfaces is not set
```

```
  Default networks flagged in outgoing updates
  Default networks accepted from incoming updates
  EIGRP-IPv4 VR(examcram) Address-Family Protocol for AS(100)
    Metric weight K1=1, K2=0, K3=1, K4=0, K5=0 K6=0
    Metric rib-scale 128
    Metric version 64bit
    Soft SIA disabled
    NSF-aware route hold timer is 240
<... output omitted ...>
```

# Route Summarization

You can scale EIGRP autonomous systems by using route summarization. Generally, as the size of an EIGRP autonomous system increases, the convergence time may increase. Scaling EIGRP requires you to summarize routes hierarchically.

EIGRP summarization is done on network prefixes at the interface level. Once summarized, prefixes within the summary aggregate are suppressed, and the summary aggregate is used instead of the original prefixes. However, the summary aggregate is not advertised until a prefix matches it. Route summarization also helps during periods of convergence when a route goes active. In this case, summarization creates a query boundary and shrinks the query domain.

Cisco routers support both auto and manual summarization. With auto summarization, EIGRP summarizes networks at the major network boundaries. It is therefore recommended that auto summarization be turned off and that you handle summarization manually. Cisco IOS Version 15 and later have auto summarization disabled by default. If needed, however, you can turn off auto summarization under the EIGRP process by using the command **no auto-summary**. Manual route summarization is done in classic mode by using the **ip summary-address eigrp** command in interface configuration mode and in named mode by using the **summary-address** command.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. What protocol number does EIGRP use to identify packets?
   - ❍ **A.** 87
   - ❍ **B.** 88
   - ❍ **C.** 89
   - ❍ **D.** 90

2. What is the default hello timer value that EIGRP uses on high-speed interfaces?
   - ❍ **A.** 3 seconds
   - ❍ **B.** 5 seconds
   - ❍ **C.** 10 seconds
   - ❍ **D.** 15 seconds

## Answers

1. **B** is correct. EIGRP uses protocol number 88 to identify packets.
2. **B** is correct. On high-speed interfaces, EIGRP uses a hello timer of 5 seconds.

# Open Shortest Path First (OSPF)

This section looks at the fundamentals of OSPF and how communication happens between OSPF routers. It also explores basic OSPF configuration and optimizations as well as advanced OSPF topics like path selection, route summarization, and OSPFv3.

Open Shortest Path First (OSPF) is an industry-standard interior gateway protocol (IGP) that uses a link-state algorithm. It was developed by the OSPF Working Group of the Internet Engineering Task Force (IETF). OSPF was explicitly designed for IP networks. OSPF supports plaintext and MD5 authentication, variable-length subnet masking (VLSM), and classless routing. Currently, there are two versions of OSPF in use:

- **OSPFv2:** This version, which was initially defined in RFC 2328, supports IPv4.
- **OSPFv3:** This version, which was defined in RFC 5340, modifies the original structure to support IPv6.

OSPF uses the shortest path first algorithm to build and calculate the shortest path to every known network destination. The shortest path calculation is done using the Dijkstra shortest path first algorithm. The algorithm works like this:

1. When a router running OSPF initializes or changes routing information, a router generates a link-state advertisement. The advertisement represents the collection of all link states on that router.
2. All OSPF routers exchange link states by means of flooding. Each router that receives a link-state update stores a copy in its link-state database, and the update is then propagated to other OSPF routers.
3. After each router's database is updated, the router calculates the shortest path tree to all network destinations by using the Dijkstra shortest path first algorithm. The network destinations, their costs, and the next hops to reach those network destinations form the IP routing table.
4. Although OSPF does not refresh its routing updates periodically, it refloods link-state advertisements (LSAs) every 30 minutes. When an update reaches a life span of 60 minutes, it is removed from the link-state database (LSDB), and the router performs a new shortest path first calculation. In addition, the router floods the LSA to other routers so that they can remove the LSA as well. This update, called a *paranoid update*, is used to refresh the LSDB; the paranoid update time is 30 minutes. Thus, every

LSA, regardless of changes, gets re-flooded at least once every 30 minutes by default. Type 1 and type 2 LSAs trigger a Dijkstra shortest path first algorithm calculation. LSA types are covered later in this section.

When there are no changes in the OSPF network, such as to change the cost of a link or add or delete a network, OSPF is relatively quiet.

# OSPF Cost

The cost, or metric, of an OSPF interface indicates the overhead required to send packets across a certain interface. The cost is inversely proportional to the bandwidth of that interface, and the higher the bandwidth of an interface, the lower the cost. The formula for calculating the cost is as follows:

Cost = 100 Mbps/Bandwidth of the Interface, in Mbps

When OSPF uses the default cost settings, there is no differentiation in the link cost that is associated with a Fast Ethernet interface and a 10 Gigabit Ethernet interface. For example, Table 2.4 shows that the Fast Ethernet, Gigabit Ethernet, and 10 Gigabit Ethernet interfaces all use the same OSPF cost of 1.

TABLE 2.4 **OSPF Interface Types and Their Default Costs**

| Interface Type | OSPF Cost |
|---|---|
| T1 | 64 |
| Ethernet | 10 |
| Fast Ethernet | 1 |
| Gigabit Ethernet | 1 |
| 10 Gigabit Ethernet | 1 |

Changing the reference bandwidth to a higher value allows you to differentiate the costs between the higher-speed interfaces. Setting this value too high causes the low-bandwidth interfaces to be indistinguishable.

When you change the reference bandwidth on one OSPF router, you need to change it on all other OSPF routers so that the shortest path first algorithm uses the same logic and prevents routing loops. You change the reference bandwidth in the OSPF process by using the command **auto-cost reference-bandwidth** *bandwidth-in-mbps*. By doing this, you change the reference bandwidth for all of the OSPF interfaces that are associated with that process. You can change the cost on an interface manually by using the command **ip ospf cost** *1-65535*.

# OSPF Authentication

You can authenticate OSPF packets so that a router will only accept and process OSPF packets that are successfully authenticated. This ensures that rogue OSPF packets do not jeopardize the integrity of the OSPF domain, as they will be ignored. By default, an OSPF router uses null authentication, which basically means OSPF packets are not authenticated. For OSPF, you can use a simple password or MD5 for authentication:

- **Simple password authentication:** A simple password is passed in plaintext inside the OSPF packet for authentication. A cybercriminal who captures such a packet knows the password and can compromise the integrity of the OSPF routing domain. Therefore, you should avoid simple password authentication in OSPF.
- **MD5 authentication:** As with EIGRP MD5 authentication, the MD5 keyed digest in each OSPF packet prevents the introduction of unauthorized routing messages from unapproved sources. The MD5 digest is included in each OSPF packet instead of a password. This means it is almost impossible for a cybercriminal to determine the actual authentication details and compromise the integrity of the routing domain. Therefore, you should use MD5 in a production environment.

To set up MD5 authentication, you need to define the key ID and key under the OSPF interface and then enable it. You create the key by using the command **ip ospf message-digest-key** *key-id* **md5** *key*. You then enable authentication on the interface by using the command **ip ospf authentication message-digest**.

You can also enable configuration for all OSPF interfaces in an entire area. You do this by using the command **area** *area-id* **authentication message-digest**. You need to define the key ID and key by using the command **ip ospf message-digest-key** *key-id* **md5** *key* under the interface.

You can also use cryptographic authentication, which allows you to configure a keychain on the OSPF interface to authenticate OSPF packets using HMAC and SHA algorithms. You can define a new keychain or use a keychain that is defined already and being used by another protocol.

Example 2.9 shows the process of defining a keychain and lifetime to be used for OSPF cryptographic authentication.

EXAMPLE 2.9 **Defining a Keychain**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# key chain ExamCram1
R1(config-keychain)# key 1
R1(config-keychain-key)# key-string ThisIsASampleKey123
R1(config-keychain-key)# cryptographic-algorithm hmac-sha-256
R1(config-keychain-key)# send-lifetime local 10:00:00 25 October 2021
infinite
R1(config-keychain-key)# end
R1#
```

Example 2.10 shows how to define authentication on an interface by using the previously created keychain and the **ip ospf authentication key-chain** *key-chain* command. The key-chain name in this example is **ExamCram1**. command.

EXAMPLE 2.10 **Defining Cryptographic Authentication Using a Keychain**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# interface GigabitEthernet 0/0
R1(config-if)# ip ospf authentication key-chain ExamCram1
R1(config-if)# end
R1#
```

# OSPF Areas

OSPF provides overall scalability for an entire OSPF routing domain by splitting the OSPF routing domain into multiple OSPF areas. An OSPF area serves as a logical grouping of router interfaces. The membership for an area is set at the interface level, and the area ID is propagated in OSPF hello packets. An interface can belong to only one area. All of the routers within a particular area maintain an identical copy of the link-state database (LSDB). As you introduce more network links and routers within an area, the size of the OSPF area grows. You may need to include multiple areas in your designs to improve the overall performance of the OSPF calculations and scale up your routing domain.

Using a single OSPF area simplifies a network topology but has several disadvantages:

- A full shortest path first calculation needs to run on every router when a link flaps in an area.
- Within a single area, the LSDB size increases and can become unmanageable for the router due to memory or processing issues.
- As the LSDB for a single area grows in size, it consumes more memory and takes longer to handle the shortest path first computation.
- No summarization of route information can occur due to the strict rules, principles, and guidelines required by OSPF within a single area to ensure that all routers within that particular area maintain an identical copy of the LSDB.

For a multi-area OSPF routing domain, one area must be labeled as Area 0. Area 0 is also referred to as the backbone, and all additional areas need to be connected to the backbone. To better visualize this, think of a hub and spoke topology. In such a topology, the hub is Area 0, and all other areas are the spokes that are directly connected to Area 0.

OSPF expects all areas to inject routing information into the backbone, and then the backbone will disseminate the information to other areas. If you cannot have direct physical connectivity to the backbone, then you need to configure a virtual link. A virtual link provides a disconnected area with a logical path to the backbone. However, to ensure maximum performance and scalability, you should do your best to avoid virtual links and try to directly connect all spokes to the hub. You should use virtual links only as a last resort and only temporarily until you find a way to make a direct connection.

Virtual links serve two purposes:

- **Linking an area that does not have a physical connection to the backbone:** For example, say that you have Area 0 connected to Area 113, but you need to add another area (Area 51), and the only way to do so is to connect it to Area 113, as you can't get Area 51 directly connected to Area 0. Eventually you will find a way to connect Area 51 directly to Area 0.
- **Patching the backbone in the event that a discontinuity of Area 0 occurs:** You would do this, for example, when two companies merge and need to connect their Area 0s but can't directly connect them yet. Eventually, you will find a way to directly connect them so they are one large Area 0 instead of two smaller ones.

# Neighbors and Adjacencies

OSPF routers that share a common segment become neighbors on that segment. Routers establish neighborships on a segment by periodically exchanging hello packets out interfaces using the IP multicast destination address 224.0.0.5.

> **ExamAlert**
>
> For the ENCOR exam, it is important to know the parameters that routers must agree on before becoming OSPF neighbors.

To become OSPF neighbors, two routers must agree on the following:

- **Area ID:** When two routers are on a common segment, their interfaces must belong to the same area on that segment.
- **Authentication:** Routers trying to become neighbors must exchange the same password/MD5 hash on a particular segment.
- **Hello and dead intervals:** Hello packets are used to form neighborships and act as keepalives on each segment to maintain neighborships. The hello interval is the number of seconds that hello packets are sent out an interface. The dead interval is the number of seconds without hello packets after which the router declares the OSPF router down. In order to form a neighborship, the OSPF hello and dead timers must match on the interfaces that are forming a neighborship with each other.
- **Stub area flag:** Two routers need to agree on the stub area flag in the hello packet before they can become neighbors.
- **Router ID:** The router IDs must be unique for neighborships to form. Ideally, you should statically configure an IP address on a loopback interface that is always up for the router ID.
- **MTU:** The interfaces that are forming a neighborship must have the same MTU configured. If there is an MTU mismatch, the routers will get stuck in the Exstart/Exchange state.

After the neighboring process, adjacency is the next step. Adjacent routers move past the hello phase and proceed to database exchange. In multi-access environments such as Ethernet, before moving on to the database exchange, OSPF must elect a DR and a backup DR (BDR) on the segment during the two-way stage to minimize the amount of information exchange on a segment during the database exchange. This gives routers on the segment participating in OSPF a central point of contact for information exchange, which cuts down

on information exchange as each router does not need to communicate with every other router. With DR and BDR election, the routers with the highest OSPF priority become the DR and BDR, respectively. In the event of a tie, the router with the highest router ID wins. All other routers are considered DROTHER. On point-to-point interfaces, there is no concept of DR and BDR.

For inter-router communication, OSPF operates directly over IPv4, using its own protocol number, 89. It uses multicast where necessary to reduce traffic. These are the multicast addresses that an OSPF router uses:

- **AllSPFRouters:** IPv4 address 224.0.0.5. All routers running OSPF should be able to receive these packets.
- **AllDRouters:** IPv4 address 224.0.0.6. Communication with DRs uses this address.

ExamAlert

For the ENCOR exam, it is important to know the states that an interface goes through before forming a fully adjacency.

Let us review the states an interface passes through before forming a full adjacency:

- **Down:** No hellos have been received from any router on the segment, but hellos are being sent.
- **Attempt:** On non-broadcast multi-access links, this state indicates that no recent hello has been received from the neighbor. This state is only seen when forming static neighborships on segments that don't support multicasting.
- **Init:** A hello packet was detected on the interface, but bidirectional communication has not yet been established. At this point on the interface, you are successfully sending and receiving hellos. This state is only seen when forming dynamic neighborships on segments that support multicasting.
- **Two-way:** Bidirectional communication with the neighbor has occurred, and the router saw its own router ID in the neighbor field of the hello packet coming from the neighbor. At the end of this stage, DR and BDR elections occur, and routers decide whether to proceed with the database exchange.

- **Exstart:** Routers establish the initial sequence number that is to be used for information exchange. Sequence numbers ensure that routers have the most recent information. This is also the stage where the routers determine who will be in control of the conversation so that the database exchange is smooth. The router with the highest router ID is the winner.
- **Exchange:** Routers send database descriptor packets describing their LSDB. If you see any of your routers stuck in the Exstart or Exchange state, there is likely an MTU mismatch between the two interfaces forming the adjacency. Routers do not by default move beyond this state if the MTUs are not matching. This ensures that mismatched MTUs don't cause fragmentation of OSPF messages and therefore potentially damage or corrupt OSPF packets.
- **Loading:** Routers finalize the information exchange. At this stage, routers build a link-state request list and a link-state retransmission list.
- **Full:** Adjacency is completed at this stage. The neighboring routers are fully adjacent and have synchronized copies of the LSDB.

# OSPF Packet Types

Table 2.5 describes the packet types that the OSPF uses.

TABLE 2.5 **OSPF Packet Types**

| Type | Packet Name | Description |
|---|---|---|
| 1 | Hello | Used for discovering and maintaining neighbors. Hello is used in the Init, Attempt, and 2-Way states. |
| 2 | Database descriptor (DBD) | Used for summarizing link-state database contents. DBDs are used as part of the Exstart and Exchange states. |
| 3 | Link-state request (LSR) | Used to request link-state database updates. LSRs are used in the Loading state. |
| 4 | Link-state update (LSU) | Used to send link-state database updates. LSUs are used in the Loading state. |
| 5 | Link-state ack (LSACK) | LSACKs are used to ensure reliable transmission of LSAs. |

# Basic OSPF Configuration

You can configure the OSPF process and use network statements to identify the interfaces that it will use and the areas in which those interfaces should

participate. You do this by using the command **network** *ip-address wildcard-mask* **area** *area-id*. You can also configure OSPF specifically on an interface by using the command **ip ospf** *process-id* **area** *area-id*. You can verify that OSPF is running on the correct interface after configuration by using the command **show ip ospf interface** [**brief** | *interface-id*]. The command **show ip ospf neighbor [detail]** allows you to verify the OSPF neighbor table. You verify routes learned through OSPF by using the command **show ip route ospf**.

OSPF configuration is straightforward, so this section does not cover it. However, this section does provide some verification examples.

Example 2.11 shows the output of the **show ip ospf interface** and **show ip ospf interface brief** commands. The **show ip ospf interface** command shows information such as the OSPF interface's address, DR and BRD IDs, timers, and neighbors that the router is adjacent with. The command **show ip ospf interface brief** provides a quick summary of information such as the OSPF interfaces and the area, IP address, cost, and state of each interface.

EXAMPLE 2.11 **Verifying OSPF Interfaces**

```
R2# show ip ospf interface
Loopback1 is up, line protocol is up
  Internet Address 2.2.2.2/32, Area 0, Attached via Network Statement
  Process ID 1, Router ID 2.2.2.2, Network Type LOOPBACK, Cost: 1
  Topology-MTID    Cost    Disabled    Shutdown      Topology Name
        0           1         no                      Base
  Loopback interface is treated as a stub Host
GigabitEthernet0/0 is up, line protocol is up
  Internet Address 172.16.1.2/24, Area 0, Attached via Network
Statement
  Process ID 1, Router ID 2.2.2.2, Network Type BROADCAST, Cost: 1
  Topology-MTID    Cost    Disabled    Shutdown      Topology Name
        0           1         no          no            Base
  Transmit Delay is 1 sec, State DR, Priority 1
  Designated Router (ID) 2.2.2.2, Interface address 172.16.1.2
  Backup Designated router (ID) 1.1.1.1, Interface address 172.16.1.1
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
    oob-resync timeout 40
    Hello due in 00:00:09
  Supports Link-local Signaling (LLS)
  Cisco NSF helper support enabled
  IETF NSF helper support enabled
  Index 1/1/1, flood queue length 0
  Next 0x0(0)/0x0(0)/0x0(0)
  Last flood scan length is 1, maximum is 1
  Last flood scan time is 1 msec, maximum is 1 msec
```

```
  Neighbor Count is 1, Adjacent neighbor count is 1
    Adjacent with neighbor 1.1.1.1  (Backup Designated Router)
  Suppress hello for 0 neighbor(s)
GigabitEthernet0/1 is up, line protocol is up
  Internet Address 192.168.1.2/24, Area 1, Attached via Network
Statement
  Process ID 1, Router ID 2.2.2.2, Network Type BROADCAST, Cost: 1
  Topology-MTID    Cost    Disabled    Shutdown      Topology Name
        0           1         no          no            Base
  Transmit Delay is 1 sec, State BDR, Priority 1
  Designated Router (ID) 3.3.3.3, Interface address 192.168.1.3
  Backup Designated router (ID) 2.2.2.2, Interface address 192.168.1.2
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
<... output omitted ...>
R2#
R2# show ip ospf interface brief
Interface    PID   Area            IP Address/Mask    Cost  State Nbrs F/C
Lo1          1     0               2.2.2.2/32         1     LOOP  0/0
Gi0/0        1     0               172.16.1.2/24      1     DR    1/1
Gi0/1        1     1               192.168.1.2/24     1     BDR   1/1
```

Example 2.12 shows the output of the **show ip ospf** neighbor command. This output shows the neighbors, state, hold time, address, and the interface where neighborship was established.

EXAMPLE 2.12 **Verifying OSPF Neighbors**

```
R2# show ip ospf neighbor

Neighbor ID  Pri  State      Dead Time  Address       Interface
1.1.1.1        1  FULL/BDR   00:00:39   172.16.1.1    GigabitEthernet0/0
3.3.3.3        1  FULL/DR    00:00:39   192.168.1.3   GigabitEthernet0/1
R2#
```

Example 2.13 shows the routes learned through OSPF. As you can see from the output, O indicates that these are routes learned via OSPF.

EXAMPLE 2.13 **Verifying OSPF Routes**

```
R2# show ip route ospf
Codes: L - local, C - connected, S - static, 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
       i - IS-IS, su - IS-IS summary, 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, H - NHRP,
l - LISP
       a - application route
       + - replicated route, % - next hop override, p - overrides from
PfR

Gateway of last resort is not set

      1.0.0.0/32 is subnetted, 1 subnets
O        1.1.1.1 [110/2] via 172.16.1.1, 00:12:04, GigabitEthernet0/0
      3.0.0.0/32 is subnetted, 1 subnets
O        3.3.3.3 [110/2] via 192.168.1.3, 00:11:11, GigabitEthernet0/1
R2#
```

Example 2.14 shows the output of the **show ip protocol** command, which provides information about the routing process and the role that the router is performing. In this case, this is an ABR.

EXAMPLE 2.14 **Verifying Protocols**

```
R2# show ip protocol
*** IP Routing is NSF aware ***
<... output omitted ...>

Routing Protocol is "ospf 1"
  Outgoing update filter list for all interfaces is not set
  Incoming update filter list for all interfaces is not set
  Router ID 2.2.2.2
  It is an area border router
  Number of areas in this router is 2. 2 normal 0 stub 0 nssa
  Maximum path: 4
  Routing for Networks:
    2.2.2.2 0.0.0.0 area 0
    172.16.1.0 0.0.0.255 area 0
    192.168.1.0 0.0.0.255 area 1
  Routing Information Sources:
    Gateway         Distance      Last Update
    3.3.3.3              110      00:11:29
    1.1.1.1              110      00:12:22
  Distance: (default is 110)

R2#
```

# Router ID (RID)

In OSPF, the router ID (RID) is dynamically allocated and is the highest IP address of any loopback interface in the *up* state. If there are no loopback interfaces in the *up* state, then the highest IP address of any active physical interface in the *up* state is used when the OSPF process initializes.

The OSPF topology is built on the RID. Configuration of a static RID helps with troubleshooting and reduces LSAs when a RID changes in an OSPF environment. An RID is four octets in length and generally is represented by an IP address on the router for simplicity. (However, this is not a requirement.)

You can statically assign an RID by using the command **router-id** *router-id* under the OSPF process. You can use the command **clear ip ospf process** to restart the OSPF process and use a new RID.

# Passive Interfaces

A passive interface in OSPF prevents the sending of OSPF hellos and prevents the processing of any received OSPF packets. Making an interface passive helps with hardening the routing protocol and reduces the use of resources. It prevents someone from plugging in an unauthorized OSPF router on an OSPF-enabled network segment and introducing false routes into the network. A passive interface adds the network segment to the LSDB, but it prohibits the formation of adjacencies.

You use the command **passive** *interface-id* under the OSPF process to make an interface passive, and you use the command **passive interface default** to make all interfaces passive. You can then use the command **no passive** *interface-id* to allow a particular interface to process OSPF packets.

# Default Route Advertisements

You can advertise a default route into an OSPF domain. This is useful when you are routing from an OSPF domain toward external networks or toward the Internet. You advertise the default route by using the command **default-information originate [always] [metric** *metric value*] **metric-type** *type-value*] in the OSPF process. The **always** option installs a default route even if one does not exist in the RIB. You can change the metric with the **metric** option and change the metric type with the **metric-type** option.

# OSPF Optimizations

You can tune OSPF in a number of ways. You can adjust the reference bandwidth to differentiate between slower and faster links (such as between 1 Gbps and 10 Gbps, as illustrated earlier in this chapter). You can use the **auto-cost reference-bandwidth** *bandwidth-in-mbps* command to change the cost in the OSPF process for all interfaces. For optimization, you can also set the cost manually on an interface by using the command **ip ospf cost** *1-65535*.

You can also fine-tune the OSPF hello and dead timers. You adjust the hello timer by using the command **ip ospf hello-interval** *1-65535* in interface configuration mode. You adjust the dead timer by using the command **ip ospf dead-interval** *1-65535* in interface configuration mode. You can verify the timers by using the command **show ip ospf interface**. The hello and dead timers must match for all routers in a segment in order for those routers to become neighbors. It is therefore important that the timers be changed at the other end.

You can influence the DR placement in a network. You can raise the priority above the default value (which is 1) to make an interface more favorable than other interfaces that have the default value. You change the priority by using the interface configuration command **ip ospf priority** *0-255*.

# Link-State Advertisements (LSAs)

As OSPF routers become adjacent, the LSDB is synchronized between routers by exchanging LSAs within LSUs. As directly connected links are added to or removed from a router's database, LSAs must be flooded out all active OSPF interfaces. OSPF uses six LSA types for IPv4 routing, and LSA types 1, 2, and 3 are used for building the shortest path first tree for intra-area and interarea routes. Specifically, LSA type 1 and type 2 are used to build the shortest path first tree for an intra-area router, LSA type 3 and type 4 are used for interarea routes, and LSA type 5 and type 7 are used for external routes.

**ExamAlert**

It is important that you commit to memory the different LSA types and their uses before taking the exam.

Let us review the LSA types:

- **Type 1 (router LSA):** Advertises the LSAs that are originated within an area. It is generated by all routers in an area and only flooded in that same area.
- **Type 2 (network LSA):** Advertises the multi-access network segment that is attached to a DR within an area. It is generated by the DR of a multi-access segment and only flooded in that same area.
- **Type 3 (summary LSA):** Advertises the network prefixes that have originated from a different area. It is generated by ABRs and only flooded in the directly connected area.
- **Type 4 (ASBR summary):** Advertises the ID of the ABR that needs to be used to reach an ASBR if the ASBR is not in the area where the type 5 LSA is being seen. This LSA is generated by the ABR and flooded into the directly connected area the ASBR is not in to help the routers in the area get to the ASBR advertising the type 5 LSA.
- **Type 5 (AS external LSA):** Advertises LSAs for routes that were redistributed and therefore external to the OSPF routing domain. ASBRs generate these LSAs, and they are flooded, unmodified, throughout the entire OSPF routing domain to all routers in all areas (hence the needs for type 4 LSAs).
- **Type 7 (NSSA external LSA):** Advertises redistributed routes in not-so-stubby areas (NSSAs) and is therefore external to the OSPF routing domain. ASBRs generate these LSAs, and they are only flooded into the NSSA the ASBR is part of. Any ABR connected to the NSSA will convert the type 7 LSA into a type 5 LSA as it is being sent into Area 0. Type 7 LSAs exist only in the NSSA and nowhere else.

# OSPF Path Selection

OSPF uses the shortest path first algorithm to create a loop-free topology of shortest paths. All the routers use the same logic to do the shortest path calculation. Path selection prioritizes paths based on the following logic:

- **Intra-area:** The routes advertised through a type 1 LSA are always preferred over type 3 and type 5 LSAs. When multiple intra-area routes exist, the path with the lowest total path metric is installed in the RIB. If there is a tie, both routes are installed in the RIB. These routes are flagged as O in the routing table.

- **Interarea:** The next priority in selecting routes is the path with the lowest total path metric to a network destination. If there is a tie, both routes are installed in the RIB. All of the interarea paths for a route must go through Area 0. Interarea summary routes are generated by type 3 LSA on ABRs. These routes have priority over intra-area routes. If there are no type 1 LSAs, then type 3 is chosen. These routes are flagged as O IA in the routing table.
- **External type 1:** External type 1 LSAs use the redistribution metric in addition to the lowest path metric to get to the ASBR that advertises the route. Type 1 path metrics are lower for routes closer to the originating ASBR. A type 1 cost includes an additional external cost with the internal cost that is used to reach that route. A type 1 route is always preferred over a type 2 route for the same destination. If there are no type 3 LSAs, E1 and N1 routes are not installed in the RIB simultaneously, and N1 is given preference for a typical NSSA, which prevents E1 from being installed on the ABR.
- **External type 2:** External type 2 routes do not increment in metric, regardless of the path metric to the ASBR. An ABR does not install E2 and N2 routes in the RIB at the same time. Preference is given to N2 routes for a typical NSSA, preventing E2 routes from being installed on the ABR.

External type 1 and type 2 routes are flagged as O E1 and O E2 in the routing table. NSSA external type 1 and NSSA external type 2 are flagged as O N1 and O N2 in the routing table.

Multiple routes to the same destination are preferred in the following order:

1. Intra-area, O
2. Interarea, O IA
3. External E1, O E1
4. External E2, O E2

Whenever OSPF identifies multiple paths to a network destination, the routes are installed in the routing table using equal-cost multipathing (ECMP). The default maximum number of paths is four, and the command **maximum-paths** *maximum-paths* is used under the OSPF process to change this default setting.

# Route Summarization

An OSPF routing domain can be split up into multiple areas to reduce the size of each LSDB. Although the routers and networks remain the same within the routing domain, the detailed type 1 and type 2 LSAs are exchanged for simpler type 3 LSAs between areas. Therefore, all routers in OSPF have an identical copy of the LSDB, and an OSPF area needs to be able to cater to routers with both low and high CPU resources. Route summarization speeds up shortest path first calculation. Because all routers within an area must have identical copies of the LSDB, summarization can only happen between areas on ABRs or as routes are being redistributed into OSPF on an ASBR. Interarea summarization (that is, LSA type 3 summarization) is done using the command **area** *area-id* **range** *network subnet-mask* [**advertise** | **not-advertise**] [**cost** *metric*] under the OSPF process.

External summarization (that is, LSA type 5 summarization) reduces the number of external LSAs in an OSPF routing domain. With OSPF redistribution, the external routes are advertised in the OSPF routing domain as type 5 or type 7 LSAs (for the NSSA). To configure external summarization, you use the command **summary-address** *network subnet-mask* under the OSPF process.

# OSPFv3

OSPFv3, which is the latest version of the OSPF protocol, was designed for OSPF support. However, its implementation supports both the IPv4 and IPv6 address families. The protocol mechanisms covered so far in this chapter remain the same for OSPFv3. This section reviews OSPFv3 support for routing.

OSPFv3 is designed as a routing protocol for IPv4 and IPv6. The interface information includes the IPv6 prefix of the interface, the network mask, the type of network it is connected to, the devices connected to that network, and so on.

**ExamAlert**

Remembering the high-level differences between OSPFv2 and OSPFv3 would be of value for the ENCOR exam.

Before getting into our review of how OSPFv3 works, let us briefly look at some of the high-level differences between OSPFv2 and OSPFv3:

- OSPFv3 expands on OSPFv2 to provide support for IPv6 routing prefixes and the larger size of IPv6 addresses.
- The OSPFv3 routing process does not need to be explicitly created. Once you enable OSPFv3 on an interface, it causes a routing process and its associated configuration to be created.
- OSPFv3 specifies that each interface must be enabled using commands in interface configuration mode. This feature is different from OSPFv2, in which interfaces are indirectly enabled using the network command.
- With IPv6, you can configure many address prefixes on an interface. In OSPFv3, all address prefixes on an interface are included by default. You cannot select only some address prefixes to be imported into OSPFv3. Either all address prefixes on an interface are imported or no address prefix on an interface is imported.
- Unlike with OSPFv2, multiple instances of OSPFv3 can run on a link.

The following LSAs are used in OSPFv3:

- **Type 1 (router LSAs):** These LSAs describe the link state and cost of a router's links to an area. An LSA indicates whether a router is an ABR or ASBR. In OSPFv3, these LSAs have no address information and are network protocol independent. Also, in OSPFv3, the router interface information may be spread across multiple router LSAs. Receivers must concatenate all router's LSAs that were originated from a given router when running a shortest path first calculation.
- **Type 2 (network LSAs):** These LSAs describe the link-state and cost information for all routers attached to the network. Only the designated router tracks this information and can generate a network LSA. In OSPFv3, network LSAs have no address information and are network protocol independent.
- **Type 3 (interarea prefix LSAs for ABRs):** These LSAs advertise internal networks to routers in other areas (that is, interarea routes). ABRs generate summary LSAs. In OSPFv3, addresses for these LSAs are expressed using the prefix, prefix length instead of address, and mask. The default route is expressed as a prefix with a length of 0.

- **Type 4 (interarea router LSAs for ASBRs):** These LSAs advertise the location of an ASBR. Routers that are trying to reach external networks use these advertisements to figure out the best path to the next hop.
- **Type 5 (autonomous system external LSAs):** These LSAs redistribute routes from another autonomous system (usually from a different routing protocol) into OSPFv3.
- **Type 8 (link LSAs):** These LSAs have local-link flooding scope and are never flooded beyond the link with which they are associated. Link LSAs provide the link-local address of the router to all other routers that are attached to the link and inform other routers attached to the link of a list of prefixes to associate with the link. This allows a router to associate a collection of option bits with the network LSA that will be originated for the link.
- **Type 9 (intra-area prefix LSAs):** A router can originate multiple intra-area prefix LSAs for each router or transit network, each with a unique link-state ID. The link-state ID for each intra-area prefix LSA describes its association to either the router LSA or the network LSA and contains prefixes for stub and transit networks.

OSPFv3 packets use protocol ID 89, and routers communicate with each other using the local interface's IPv6 link-local address as the source. The OSPFv3 LSDB creates a shortest path tree topology based on links instead of networks. Based on the packet type, the destination address is either a unicast link-local address or a multicast link-local scoped address:

- **FF02::05:** OSPFv3 AllSPFRouters
- **FF02::06:** OSPFv3 AllDRouters designated router (DR) router

ExamAlert

For the ENCOR exam, you should know the high-level steps for OSPFv3 and IPv6 address-family configuration and verification.

The high-level steps for configuring and verifying OSPFv3 are as follows:

1. Enable IPv6 unicast routing by using the command **ipv6 unicast-routing**.
2. Configure the OSPFv3 process by using the command **router ospfv3** [*process-id*].

3. Define the router ID by using the command **router-id** *router-id*.
4. Enable OSPFv3 on an interface and assign the interface to an area by using the command **ospfv3** *process-id* **ipv6 area** *area-id*.
5. Verify neighborship by using the command **show ospfv3 ipv6 neighbor**; verify the OSPFv3-enabled interface status by using the command **show ospfv3 interface** [*interface-id*]; and show the OSPFv3 IPv6 routing table by using the command **show ipv6 route ospf**.

The high-level steps for configuring and verifying an IPv6 address family in OSPFv3 are as follows:

1. Enable IPv6 unicast routing by using the command **ipv6 unicast-routing**.
2. Configure the OSPFv3 process by using the command **router ospfv3** [*process-id*].
3. Define the router ID by using the command **router-id** *router-id*.
4. Initialize the address family by using the command **address-family** [**ipv6** | **ipv4**] **unicast**.
5. Enable OSPFv3 by using the command **area** *area-id* **range** *ipv6-prefix/prefix-length*.
6. Verify neighborship by using the command **show ospfv3 ipv6 neighbor**; verify the OSPFv3-enabled interface status by using the command **show ospfv3 interface** [*interface-id*]; and show the OSPFv3 IPv6 routing table by using the command **show ipv6 route ospf**.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. How many packet types does OSPF use for communication?
   - ❍ **A.** Two
   - ❍ **B.** Three
   - ❍ **C.** Four
   - ❍ **D.** Five

2. With OSPF, which factor is used for calculating the cost of an interface?
   - ❍ **A.** Bandwidth
   - ❍ **B.** Delay
   - ❍ **C.** Load
   - ❍ **D.** Reliability

# Answers

1. **D** is correct. OSPF uses five packet types for communication: hello, DBD, LSR, LSU, and ACK.
2. **A** is correct. By default, bandwidth is used to calculate the cost in OSPF.

# Review Questions

1. Which of the following is a pure distance vector routing protocol?
   - ❍ **A.** RIP
   - ❍ **B.** EIGRP
   - ❍ **C.** OSPF
   - ❍ **D.** BGP

2. What is the default hello timer that EIGRP uses on slower interfaces?
   - ❍ **A.** 10 seconds
   - ❍ **B.** 15 seconds
   - ❍ **C.** 60 seconds
   - ❍ **D.** 180 seconds

3. In OSPF, what is the AllSPFRouters multicast destination?
   - ❍ **A.** 224.0.0.5
   - ❍ **B.** 224.0.0.6
   - ❍ **C.** 224.0.0.8
   - ❍ **D.** 224.0.0.10

4. What is the multicast destination address that OSPFv3 uses to communicate with the AllDRouters designated routers?
   - ❍ **A.** 224.0.0.5
   - ❍ **B.** 224.0.0.6
   - ❍ **C.** FF02::05
   - ❍ **D.** FF02::06

5. True or false: Unlike OSPFv2, OSPFv3 supports running multiple instances of OSPFv3 for each link.
   - ❍ **A.** True
   - ❍ **B.** False

# Answers to Review Questions

1. **A** is correct. RIP is a pure distance vector routing protocol. With distance vector, or Bellman–Ford, algorithms, each router sends all or some portion of its routing table only to its directly connected devices.
2. **C** is correct. On slower interfaces, EIGRP uses a hello timer of 60 seconds.
3. **A** is correct. The AllSPFRouters multicast destination is 224.0.0.5.
4. **D** is correct. The AllDRouters multicast destination in OSPFv3 is FF02::06.
5. **A** is correct. OSPFv3 supports the running of multiple instances of OSPFv3 on a single link.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*
- *CCNP Enterprise Advanced Routing ENARSI 300-410 Official Cert Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers Border Gateway Protocol (BGP).

*This page intentionally left blank*

CHAPTER 3

# Understanding Layer 3: BGP

**This chapter covers the following official ENCOR 350-401 exam objective:**

- 3.2 Layer 3
- 3.2c Configure and verify eBGP between directly connected neighbors (best path selection algorithm and neighbor relationships)

Border Gateway Protocol (BGP) is a routing protocol defined in RFC 4271 for exchanging routes between autonomous systems (AS). It is the routing protocol preferred for large-scale routing and is the underlying routing protocol of the Internet. This chapter starts by covering the fundamentals of BGP. It examines the various tables that BGP uses in its operation, the various BGP message types, and neighbor states. It also examines BGP path selection and the various well-known and optional attributes. Finally, this chapter wraps up by reviewing the concept of interdomain routing through configuration and verification of a single-homed external Border Gateway Protocol (eBGP) connection.

**This chapter covers the following technology topics:**

- BGP Fundamentals
- BGP Configuration and Verification

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. What BGP message is responsible for advertising, updating, and withdrawing previously advertised routes?
2. In which BGP state are BGP sessions started and do BGP neighbors begin exchanging routes with Update messages?

### Answers

1. Update
2. Established

# BGP Fundamentals

BGP is an Internet Engineering Task Force (IETF) standard and one of the most scalable routing protocols. BGP is used as the routing protocol of the Internet and in large-scale data centers and service provider private networks. It is considered an interdomain routing protocol that is capable of providing large-scale loop-free routing between routing domains.

Originally, the purpose of BGP was to carry Internet reachability information, but it has now been expanded to carry routes for other protocols. RFC 2858 introduced multiprotocol extensions to facilitate BGP carrying routing information for IP multicast routes and multiple Layer 3 protocol address families, including IP Version 4 (IPv4), IP Version 6 (IPv6), Virtual Private Networks Version 4 (VPNv4), Connectionless Network Services (CLNS), and Layer 2 VPNs (L2VPNs). The Cisco implementation of BGP Version 4 includes support for 4-byte autonomous system numbers (ASNs) and multiprotocol extensions.

BGP is designed to run over a reliable transport protocol, and it uses TCP (port 179) as its transport protocol. As a routing protocol, BGP is primarily used to connect a local network to an external network to gain access to the Internet or connect to other organizations. BGP is also used with Multiprotocol Label Switching (MPLS) Layer 3 VPNs (L3VPNs) to interconnect customer sites so that they can communicate privately over a shared medium. When used for connecting to an external organization, external BGP (eBGP)

peering sessions are created. While BGP is referred to as an exterior gateway protocol (EGP), sometimes it is required to connect routers inside a network so they can pass BGP network layer reachability information (NLRI) with each other. Therefore, BGP peering in an organization exchanges routing information through internal BGP (iBGP) peering sessions.

BGP sees an AS as a group of routers under an organization's control, and it uses one or more IGPs to route within the AS. Using an IGP within an AS is not necessary; you could use BGP as the only routing protocol. However, using BGP as you would an IGP is not recommended as the protocol was not designed to route within an AS. BGP does not support dynamic neighbor discovery, and it has a complex path selection process and long convergence delays. A large organization that requires connectivity to the Internet must obtain an ASN. An ASN is a 4-byte (32-bit) value; 4,294,967,295 unique ASNs are possible. This extended version expands from the original 2-byte (16-bit) value that provided 65,535 ASNs.

**ExamAlert**

For the ENCOR exam, make sure you completely understand the various tables that are used by BGP.

There are two blocks of private ASNs that organizations can use as long as they are not exchanged publicly on the Internet. These are as follows:

- 64,512–65,535 within the 16-bit ASN range
- 4,200,000,000–4,294,967,294 within the 32-bit range

BGP uses three tables to store information:

- **BGP neighbor table:** BGP needs to be explicitly configured with each neighbor before it can establish adjacency. The BGP neighbor table keeps track of each of the configured neighbors and the state of their relationships. Keepalive messages are sent periodically to check the state of a relationship.
- **BGP table:** BGP neighbors exchange routes after they establish adjacency. The routes for each network collected from neighbors are placed in the BGP forwarding database. From here, using the BGP route selection process, they are then offered to the IP routing table.
- **IP routing table:** Routers compare the BGP-offered routes and choose the best route to be installed in the routing table based on where it was

learned—that is, whether the routes were learned from an external AS (where the AD equals 20) or learned from within the AS (where the AD equals 200).

BGP uses the following message types for communication:

- **Open:** Open messages are responsible for setting up and establishing BGP adjacencies. BGP routers negotiate the capability of a session before establishing a peering session. Open message includes messages related to hold time, BGP version number, ASN, the originating routers, BGP identifier, and other optional parameters.
- **Update:** Update messages are responsible for advertising, updating, and withdrawing previously advertised routes. Update messages include the NLRI, which includes the prefix and BGP path attributes. Update messages can function as keepalives to help reduce unnecessary traffic.
- **Notification:** Notification messages are used to indicate error conditions to BGP neighbors. Such a message is sent when an error is detected with a BGP session, such as hold time expiration, change in neighbor capabilities, or a BGP session reset request. A notification message causes a BGP session to close.
- **Keepalive:** Keepalive messages are used to ensure that the BGP neighbors are still alive. Keepalives are sent at one-third the hold time agreed between two routers. Cisco routers use a default hold time of 180 seconds, so the keepalive interval is 60 seconds. If the hold time is set to 0, no keepalive is sent between the BGP neighbors.

BGP goes through a number of state changes when its routing process establishes a peering session with a peer. The first three states are related to creating a TCP session, and if the TCP session does not get created, the router tries again and loops through these three states until it is successful. Once the TCP session is up and running, an open message can be used to form the BGP peering over the TCP connection. Once the Established state is reached, the other messages can be used. These are the states:

- **Idle:** This is the initial state that the BGP routing process enters when the routing process is enabled or the device is reset. In this state, the router waits on a start event from the peering router, such as a peering confirmation from the remote peer. Once a TCP connection request is received from the remote peer, the device initiates another start event to wait for a timer before starting a TCP connection with the remote peer.

- **Connect:** This is the state in which the BGP routing process detects that a peer is trying to establish a TCP session with the local BGP speaker.
- **Active:** This is the state where the BGP process tries to establish a TCP session with a peer using the ConnectRetry timer. While the BGP routing process is in this state, start events are ignored. If the BGP routing process is reconfigured or an error occurs, the routing process releases system resources and returns to the Idle start.
- **OpenSent:** After an open message is sent from the originating router, that router waits for an open message from the other router. Basically, once the TCP connection is established, the BGP routing process sends an open message to the remote peer before transitioning to the OpenSent state. If the connection fails, the routing process transitions to the Active state.
- **OpenConfirm:** In this state, BGP waits for a keepalive or notification message. Once the router receives a neighbor keepalive message, the state is then moved to Established. If the hold timer expires or a stop event occurs, the state moves to Idle.
- **Established:** This is the state where the BGP session is established, and BGP neighbors begin exchanging routes with update messages. Once update and keepalive messages are received, the hold timer is reset. If the hold timer expires, the router receives an error notification, and the BGP process moves the neighbor back to the Idle state.

BGP uses a path vector routing algorithm to exchange network reachability information with other BGP-speaking devices. The network reachability information is exchanged in routing updates. It includes the network number, path-specific attributes, and list of autonomous system numbers that a route must transit to reach a destination network. The list of AS numbers is contained in the AS_PATH attribute.

BGP prevents routing loops by rejecting any routing update that contains the local autonomous system number because such an update indicates that the route has already passed through that autonomous system, and therefore a loop would be created. The BGP path vector routing algorithm is a combination of a distance vector routing algorithm and AS path loop detection.

By default, BGP selects a single path as the best path to a destination host or network. The best path selection algorithm analyzes the path attributes to determine which route is installed as the best path in the BGP table. Each path carries well-known mandatory, well-known discretionary, and optional transitive attributes that are used in BGP best-path analysis. Well-known attributes

must be recognized by all BGP implementations, but optional attributes do not need to be recognized by all BGP implementations. Let us look at these further:

- **Well-known mandatory:** Attributes must be included with every prefix advertisement. The well-known mandatory attributes include:
  - **Origin:** This attribute is set when a router is first originating routes in BGP.
  - **AS_PATH:** This attribute is the sequence of AS numbers through which the network is accessible.
  - **Next_Hop:** This attribute is the IP address of the next-hop router (that is, the router to which the receiving router forwards IP packets to reach a destination). The next-hop attribute is modified as the route passes through the network.
- **Well-known discretionary:** Attributes may or may not be included with the advertised prefix. The well-known discretionary attributes include the following:
  - **Local preference:** This attribute is used for achieving a consistent routing policy for traffic that exits an AS.
  - **Atomic aggregate:** This attribute is attached when a route is created as a result of aggregation. This attribute indicates that information that was present in the original routing update may have been lost due to route summarization.
- **Optional transitive:** Attributes do not have to be recognized by all implementations and can be set so that they are transitive and stay with route advertisements from AS to AS. The optional transitive attributes include:
  - **Aggregator:** This attribute is used to identify the AS and the router within that AS that created a route summarization.
  - **Community:** This attribute is a numeric value that can be attached to routes as they pass a point in the network. For example, for route selection purposes, BGP routers can examine the community value at different points.
- **Optional nontransitive:** Attributes cannot be shared from AS to AS. An optional nontransitive attribute is multi-exit discriminator (MED), which is used to influence the inbound traffic into an AS from an AS with multiple entry points.

BGP assigns the first valid path as the current best path. It then compares the best path with the next best path in the list until it reaches the end of the list of valid paths.

**ExamAlert**

For the ENCOR exam, you can use the following mnemonic to remember BGP best-path selection: "We Love Oranges As Oranges Mean Pure Refreshment" for Weight, LOCAL_PREF, Originated Locally, AS_PATH, ORIGIN Type, MED, Paths, and RID.

The following rules are used, in this order, for determining the best path:

1. Prefer the path with the highest WEIGHT (local to a router).
2. Prefer the path with the highest LOCAL_PREF (global within AS).
3. Prefer the path that was locally originated via a **network** or **aggregate** BGP subcommand or through redistribution from an IGP.
4. Prefer the path with the shortest AS path (that is, the smallest number of autonomous systems in the AS_PATH attribute).
5. Prefer the path with the lowest Origin type.
6. Prefer the path with the lowest MED. (MEDs are compared only if the first AS in the AS sequence is the same for multiple paths.)
7. Prefer eBGP over iBGP paths.
8. Prefer the path with the lowest IGP metric to the BGP next hop.
9. If both paths are external, prefer the path that was received first (that is, the oldest route).
10. Prefer the route that comes from the BGP router with the lowest router ID (RID).
11. Prefer the path from the lowest neighbor address (which is the IP address used in the BGP neighbor configuration).

To facilitate BGP load balancing over multiple paths, you use the **maximum-paths** *number-of-paths* command. BGP can perform load balancing across a maximum of six paths.

Example 3.1 looks at best path selection in BGP when there are two exit paths out of an AS. The **show ip bgp** *network* command displays the entries in the BGP table for a particular network. This example shows the **show ip bgp**

**10.0.20.0/24** command used to go over the process of best-path selection for the network 10.0.20.0/24.

EXAMPLE 3.1 **BGP Table to Demonstrate Best Path Selection**

```
R1# show ip bgp 10.0.20.0/24
BGP routing table entry for 10.0.20.0/24, version 4
Paths: (2 available, best # 2, table default, RIB-failure(17))
  Advertised to update-groups:
     1
  Refresh Epoch 1
  65001
    192.168.2.1 from 192.168.2.1 (10.0.20.2)
      Origin IGP, metric 0, localpref 100, valid, external
      rx pathid: 0, tx pathid: 0
  Refresh Epoch 1
  65001
    192.168.1.1 from 192.168.1.1 (10.0.20.1)
      Origin IGP, metric 0, localpref 100, valid, external, best
      rx pathid: 0, tx pathid: 0x0
```

BGP selects the better of these two paths by considering the various attributes covered next. In the output shown here, BGP compares the available paths and selects path 2 as the best path, based on its lower router ID.

Here is a comparison of path 1 and path 2 from Example 3.1:

- Both paths have reachable next hops.
- Both paths have a WEIGHT of 0.
- Both paths have a LOCAL_PREF of 100.
- Both paths are learned.
- Both paths have AS_PATH length 1.
- Both paths are of origin IGP.
- The paths have different neighbor AS values, so you ignore MED.
- Both paths are internal.

Based on this information, we can determine that path 2 is better than path 1 because it has a lower router ID.

Keep in mind that BGP was not intended to be a fast routing protocol. It was created to provide more administrative control over route path selection. BGP path selection involves manipulating BGP attributes. For example, you can set the default local preference by using the command **bgp default local-preference**. Or you can set the default MED for redistributed routes by using the command **default-metric** under the BGP process. You can manipulate BGP attributes by using route maps. Route maps allow you to change attributes for certain neighbors or routes only. Once you change an attribute, you must tell BGP to apply the changes. You do so by clearing the BGP session using the command **clear ip bgp *** or by using the command **clear ip bgp * soft in | out**. The first method, which is a *hard reset*, tears down the BGP session, removes BGP routes from the peer, and is the most disruptive method. The other method, a *soft reset*, invalidates the BGP cache and requests a full advertisement from its BGP peer.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which of the following BGP message types is used for indicating an error condition to a BGP neighbor?
   - ❍ **A.** Open
   - ❍ **B.** Update
   - ❍ **C.** Notification
   - ❍ **D.** Keepalive

2. Which well-known attribute is attached when a route is created as a result of aggregation?
   - ❍ **A.** Aggregate
   - ❍ **B.** Atomic aggregate
   - ❍ **C.** Community
   - ❍ **D.** Local preference

## Answers

1. **C** is correct. A notification message is used for indicating an error condition to a BGP neighbor. It is sent when an error with a BGP session is detected.
2. **B** is correct. Atomic aggregate is an attribute that is attached when a route is created as a result of aggregation. It indicates that information that was present in the original routing update may have been lost due to route summarization.

# BGP Configuration and Verification

This section reviews BGP peers' initial communication and messages before getting into BGP configuration and verification. Two routers that form a TCP connection in order to exchange BGP routing information are considered *peers*, or *neighbors*. BGP peers initially exchange full BGP tables. After this initial exchange, peers send incremental updates as the routing table changes. BGP keeps a version number for the BGP table, and this version number is the same for all the BGP peers. The version number changes when BGP updates the table with routing information changes. The sending of keepalive packets ensures that the connections between BGP peers are alive. Notification packets go out in response to errors or special conditions.

The configuration example in this section shows a simple eBGP scenario with a service provider (SP) and a customer with two routers. You will establish different eBGP sessions between the SP and customer environment, and you will simulate the advertisement of networks into the environment by advertising a loopback interface on each router. The following basic steps are involved in this simple BGP configuration:

**ExamAlert**

For the ENCOR exam, make sure you have an understanding of the configuration and verification of BGP.

1. Configure the BGP routing process and enter the router configuration mode using the command **router bgp** *autonomous-system-number*. You can specify an integer from 0 and 65534 to identify the configured device to the other BGP speakers. You can only specify a single BGP AS number on a router.
2. Add the IP address of the neighbor in the specified AS to the BGP neighbor table of the local device, using the command **neighbor** *ip-address* **remote-as** *autonomous-system-number*. By default, eBGP peering session must span a maximum of only one hop. BGP does not run over an individual interface. When you manually configure neighbors, BGP runs over TCP sessions.

3. Specify the network to advertise with the **network** statement. The meaning of the **network** statement in BGP is different than in other routing protocols. In other routing protocols, it indicates on which network interface the protocol will run. In BGP, it indicates to the local router which network prefix in the routing table (that is, connected, static, learned via a different protocol) should be injected into BGP.
4. Optionally set the router ID by using the command **bgp router-id**. Configuring the router ID resets all active BGP peering sessions.
5. Verify the entries in the BGP table by using the command **show ip bgp** [*network*] [*network-mask*].
6. Use the command **show ip bgp neighbors** to show information about the TCP and BGP connections to BGP neighbors.
7. Display the status of all BGP connections by using the command **show ip bgp summary**.

Figure 3.1 shows the reference topology used for this example.

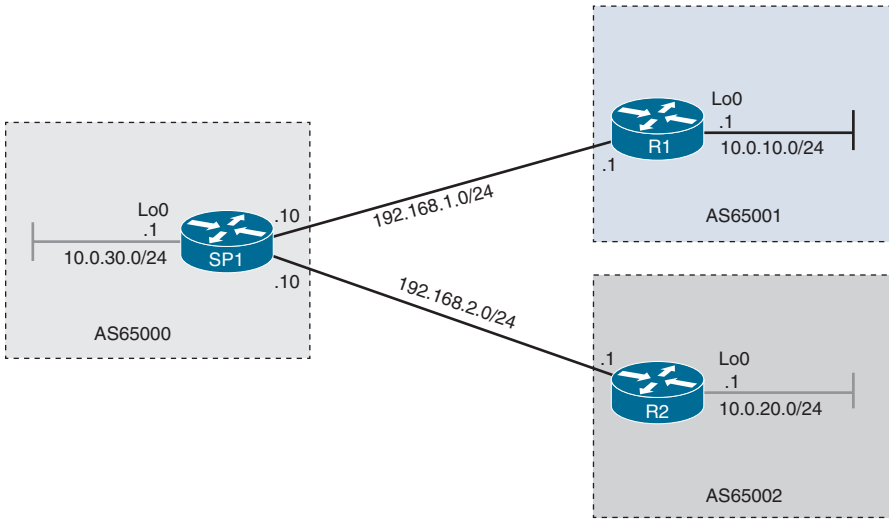



FIGURE 3.1 **BGP Configuration Topology**

Example 3.2 shows the configuration of eBGP on a service provider router (SP1) and customer routers (R1 and R2).

EXAMPLE 3.2 **Configuring eBGP on SP1, R1, and R2**

```
SP1#
SP1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SP1(config)# interface loopback0
SP1(config-if)# ip address 10.0.30.1 255.255.255.0
SP1(config-if)# exit
SP1(config)# router bgp 65000
SP1(config-router)# neighbor 192.168.1.1 remote-as 65001
SP1(config-router)# neighbor 192.168.2.1 remote-as 65002
SP1(config-router)# network 10.0.30.0 mask 255.255.255.0
SP1(config-router)# end
SP1#

R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# interface loopback0
R1(config-if)# ip address 10.0.10.1 255.255.255.0
R1(config-if)# exit
R1(config)# router bgp 65001
R1(config-router)# neighbor 192.168.1.10 remote-as 65000
R1(config-router)# network 10.0.10.0 mask 255.255.255.0
R1(config-router)# end
R1#

R2#
R2# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R2(config)# interface loopback0
R2(config-if)# ip address 10.0.20.1 255.255.255.0
R2(config-if)# exit
R2(config)# router bgp 65002
R2(config-router)# neighbor 192.168.2.10 remote-as 65000
R2(config-router)# network 10.0.20.0 mask 255.255.255.0
R2(config-router)# end
R2#
```

In Example 3.2, the BGP process starts with the command **router bgp** *autonomous-system-number*. SP1 belongs to AS 65000, R1 belongs to AS 65001, and R2 belongs to AS 65002. Then the command **neighbor** *neighbor ip-address* **remote-as** *remote-autonomous-system-number* is used to configure neighbor relationships. To advertise the network on the loopback interface, this example uses the **network** statement.

To verify the state of the BGP sessions, you use the command **show ip bgp summary**, as shown in Example 3.3.

EXAMPLE 3.3 **Verifying the State of BGP Sessions by Using the Command show ip bgp summary**

```
SP1#
SP1# show ip bgp summary
BGP router identifier 10.0.30.1, local AS number 65000
BGP table version is 4, main routing table version 4
3 network entries using 432 bytes of memory
3 path entries using 252 bytes of memory
3/3 BGP path/bestpath attribute entries using 480 bytes of memory
2 BGP AS-PATH entries using 48 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 1212 total bytes of memory
BGP activity 3/0 prefixes, 3/0 paths, scan interval 60 secs

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
192.168.1.1     4 65001       7      11        4    0    0 00:02:29        1
192.168.2.1     4 65002       7       9        4    0    0 00:02:33        1
SP1#
```

The command **show ip bgp summary** shows one line of text for each configured neighbor and its status. It shows the following:

- The IP address of the neighbor.
- The BGP version number that the router uses when communicating with the neighbor.
- The AS number of the remote neighbor.
- The number of messages and updates received from the neighbor since the session was established.
- The number of messages and updates sent from the neighbor since the session was established.
- The version number of the local BGP table.
- The number of messages waiting to be processed in the incoming queue.
- The number of messages in the outgoing queue that are to be transmitted.

- How long the neighbor has been in the current state and the state. (Lack of a state name indicates Established.)
- The number of received prefixes from the neighbor in the Established state.

If the state toggles between Idle and Active, then one of the most likely problems is an AS number misconfiguration.

The command **show ip bgp neighbors** *neighbor* shows the details of each configured neighbor. If you do not specify a neighbor, then all neighbors are provided in the output. In Example 3.4, *external link* indicates that the peering relationship is made via eBGP, and the peer is in a different AS. Notice that the BGP state is also Established. Notice also that this example supports IPv4 unicast. (Multiprotocol BGP (MP-BGP) configuration and verification are beyond the scope of the ENCOR exam.) Finally, the example shows a breakdown of the BGP message types sent and received.

EXAMPLE 3.4 **Using the Command show ip bgp neighbors for Verification**

```
SP1# show ip bgp neighbors 192.168.1.1
BGP neighbor is 192.168.1.1,  remote AS 65001, external link
  BGP version 4, remote router ID 10.0.10.1
  BGP state = Established, up for 00:03:29
  Last read 00:00:45, last write 00:00:27, hold time is 180, keepalive
interval is 60 seconds
  Neighbor sessions:
    1 active, is not multisession capable (disabled)
  Neighbor capabilities:
    Route refresh: advertised and received(new)
    Four-octets ASN Capability: advertised and received
    Address family IPv4 Unicast: advertised and received
    Enhanced Refresh Capability: advertised and received
    Multisession Capability:
    Stateful switchover support enabled: NO for session 1
  Message statistics:
    InQ depth is 0
    OutQ depth is 0

                         Sent       Rcvd
    Opens:                  1          1
    Notifications:          0          0
    Updates:                4          2
    Keepalives:             5          5
    Route Refresh:          0          0
    Total:                 12          8
```

```
  Do log neighbor state changes (via global configuration)
  Default minimum time between advertisement runs is 30 seconds
<...output omitted...>


SP1# show ip bgp neighbors 192.168.2.1
BGP neighbor is 192.168.2.1,  remote AS 65002, external link
  BGP version 4, remote router ID 10.0.20.1
  BGP state = Established, up for 00:03:55
  Last read 00:00:25, last write 00:00:45, hold time is 180, keepalive
interval is 60 seconds
  Neighbor sessions:
    1 active, is not multisession capable (disabled)
  Neighbor capabilities:
    Route refresh: advertised and received(new)
    Four-octets ASN Capability: advertised and received
    Address family IPv4 Unicast: advertised and received
    Enhanced Refresh Capability: advertised and received
    Multisession Capability:
    Stateful switchover support enabled: NO for session 1
  Message statistics:
    InQ depth is 0
    OutQ depth is 0

                         Sent       Rcvd
    Opens:                  1          1
    Notifications:          0          0
    Updates:                4          2
    Keepalives:             5          6
    Route Refresh:          0          0
    Total:                 10          9
  Do log neighbor state changes (via global configuration)
  Default minimum time between advertisement runs is 30 seconds
<...output omitted...>
```

Example 3.5 shows the output of the command **show ip bgp**, which displays the router's BGP table and allows you to verify that the router (SP1 in this case) received the routes that are indeed being advertised with the peer routers (R1 and R2 in this case). Network 10.0.30.0/24 is originated locally via the **network** command; notice that the next hop is 0.0.0.0. Network 10.0.10.0/24 was announced from 192.168.1.1 (neighbor R1). Network 10.0.20.0/24 was announced from 192.168.2.1 (neighbor R2).

EXAMPLE 3.5 **Using the Command show ip bgp for Verification**

```
SP1# show ip bgp
BGP table version is 4, local router ID is 10.0.30.1
Status codes: s suppressed, d damped, h history, * valid, > best,
i - internal,
              r RIB-failure, S Stale, m multipath, b backup-path,
f RT-Filter,
              x best-external, a additional-path, c RIB-compressed,
              t secondary path,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

     Network          Next Hop            Metric LocPrf Weight Path
 *>   10.0.10.0/24    192.168.1.1              0             0 65001 i
 *>   10.0.20.0/24    192.168.2.1              0             0 65002 i
 *>   10.0.30.0/24    0.0.0.0                  0         32768 i
```

Finally, Example 3.6 shows how to verify the BGP routes in the routing table that are coming in from the customer. As you can see in this example, the command **show ip route bgp** shows the BGP route in the routing table. The two routes in this case were learned via eBGP and so are marked with an AD of 20. The BGP MED metric value in this example is 0.

EXAMPLE 3.6 **Verifying BGP Routes**

```
SP1# show ip route bgp
Codes: L - local, C - connected, S - static, 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
       i - IS-IS, su - IS-IS summary, 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, H - NHRP,
l - LISP
       a - application route
       + - replicated route, % - next hop override, p - overrides from
PfR

Gateway of last resort is not set

      10.0.0.0/8 is variably subnetted, 4 subnets, 2 masks
B        10.0.10.0/24 [20/0] via 192.168.1.1, 02:15:13
B        10.0.20.0/24 [20/0] via 192.168.2.1, 02:15:17
SP1#
```

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. True or false: The **network** statement in BGP indicates to the local router which routes should be injected into the BGP table.
   - ❍ **A.** True
   - ❍ **B.** False

2. Which BGP verification command shows information about the TCP and BGP connection to BGP neighbors?
   - ❍ **A. show ip bgp summary**
   - ❍ **B. show ip route bgp**
   - ❍ **C. show ip bgp**
   - ❍ **D. show ip bgp neighbors**

## Answers

1. **A** is correct. The **network** statement in BGP indicates to the local router which routes should be injected into the BGP table. This is different from the use of the **network** statement with other routing protocols, where it indicates on which network interface the protocol will run.
2. **D** is correct. The **show ip bgp neighbors** command shows information about the TCP and BGP connections of the BGP neighbors to the local router.

# Review Questions

1. In which BGP state does the BGP process try to establish a TCP session with a peer by using the ConnectRetry timer?
   - ❍ **A.** Idle
   - ❍ **B.** Connect
   - ❍ **C.** Active
   - ❍ **D.** Established

2. True or false: The version number of a BGP table changes when BGP updates the table with routing information changes.
   - ❍ **A.** True
   - ❍ **B.** False

3. Which of the following techniques is the third selection criterion when making the BGP best path determination?
   - ❍ **A.** Weight
   - ❍ **B.** Origin (originated)
   - ❍ **C.** AS path
   - ❍ **D.** MED

4. True or false: AS_PATH is a well-known discretionary attribute.
   - ❍ **A.** True
   - ❍ **B.** False

# Answers to Review Questions

1. **C** is correct. In the Active state, the BGP process tries to establish a TCP session with a peer by using the ConnectRetry timer. This happens if the passive TCP session is not established during the first attempt.
2. **A** is correct. BGP keeps a version number of the BGP table, and this version number changes when BGP updates the table with routing information changes.
3. **B** is correct. The third BGP best-path selection prefers the path that was locally originated via the **network** or **aggregate** BGP subcommand or through redistribution from an IGP.
4. **B** is correct. AS_PATH is a sequence of AS numbers through which a network is accessible. It is a well-known mandatory attribute.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*
- *CCNP Enterprise Advanced Routing ENARSI 300-410 Official Cert Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers IP services.

*This page intentionally left blank*

CHAPTER 4

# IP Services

**This chapter covers the following official ENCOR 350-401 exam objectives:**

- 3.4 IP Services
- 3.4.a Describe Network Time Protocol (NTP)
- 3.4.b Configure and verify NAT/PAT
- 3.4.c Configure first-hop redundancy protocols, such as HSRP and VRRP
- 3.4.d Describe multicast protocols, such as PIM and IGMP v2/v3

This chapter looks at several technologies related to IP services. It starts by looking at the need for proper time synchronization in the network environment and the configuration and verification of Network Time Protocol (NTP). The second section looks at how to use Network Address Translation (NAT) to translate IP addresses from one domain to another. The third section looks at the configuration of first-hop redundancy protocols (FHRPs) to provide a redundant gateway for hosts. It examines the configuration and verification of Hot Standby Router Protocol (HSRP), Virtual Router Redundancy Protocol (VRRP), and Gateway Load Balancing Protocol (GLBP). The final section discusses the need for multicast and the protocols needed for its operation, including Protocol Independent Multicast (PIM) and Internet Group Management Protocol (IGMP) Versions 2 and 3.

**This chapter covers the following technology topics:**

- Network Time Protocol (NTP)
- Network Address Translation (NAT)
- First-Hop Redundancy Protocols (FHRPs)
- Multicast

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. Why would you set up a peer relationship between two NTP servers?
2. What is the default timeout value for IP NAT translations, and what is the command for changing the timeout value?
3. Which VRRP role is analogous to the HSRP active role?
4. What is one advantage that Auto-RP provides with multicast?

### Answers

1. NTP servers that are peers sync their clocks. Peers provide redundancy in the event that one of the NTP servers fails.
2. The default timeout value for IP NAT translations is 24 hours. This value can be changed with the **ip nat translation timeout** *seconds* command.
3. The VRRP master role is analogous to the HSRP active role.
4. One of the advantages of Auto-RP is that it provides load splitting when using different RPs and makes it possible to have different RPs serve different group ranges or serve each other as backups.

# Network Time Protocol (NTP)

It is important to ensure that the time is accurate on a system because processes and applications may depend on the time to tune their processes for proper network functionality. Some network devices use a software clock that resets when power is reset, whereas others use a hardware clock that does not reset when power is reset. You must have accurate time on network devices for several reasons:

- Verifying the validity of certificates based on expiration date and time
- Facilitating encryption key exchanges
- Managing passwords that need to be changed at a specific time
- Correlating security-based events across network devices and services such as router and network access control systems

- Troubleshooting network devices by identifying time-specific events in system logs

**ExamAlert**

For the ENCOR exam, you need to understand the importance of NTP, the NTP communication port, and NTP stratum levels.

NTP helps you maintain accurate time by synchronizing time among a set of distributed time servers and clients. It uses UDP, operating on port 123, as its transport protocol. Generally, an NTP server receives its time from an authoritative time source, such as a radio clock or an atomic clock. NTP uses a stratum to determine how many NTP hops away a network device is from the authoritative time source. For example, a stratum 1 NTP server has an authoritative time source (such as an atomic clock) directly attached, a stratum 2 server receives its time from a stratum 1 server, and so on.

NTP allows you to set up peer relationships between two networking devices. A peer can provide time on its own or can connect to an NTP server. An NTP service is most accurate when you point the local device and the remote peer to different NTP servers. With this setup, the local device will maintain the right time even if the NTP server fails because it will be able to use the time from the peer.

Figure 4.1 shows a network with two NTP stratum 2 servers and two switches. Switch 1 and switch 2 are NTP peers. Switch 1 uses the stratum 2 server 1, and switch 2 uses the stratum 2 server 2. If the stratum 2 server 1 fails, switch 1 maintains the correct time through its peer association with switch 2.

Before getting into the configuration steps, let us briefly look at some guidelines for configuring NTP:

- Prior to setting up a peer association, you should ensure that the clock is reliable (a client of a reliable NTP server). You use the **ntp server** *ip-address* [**prefer**] [**source** *interface-id*] command to do this.
- When setting up a peer association, you can configure some devices to point to one server and some devices to point to another server. This reduces the NTP processing load on one particular NTP server.
- In a configuration with one server, all clients should point to the NTP server and should not peer with each other.

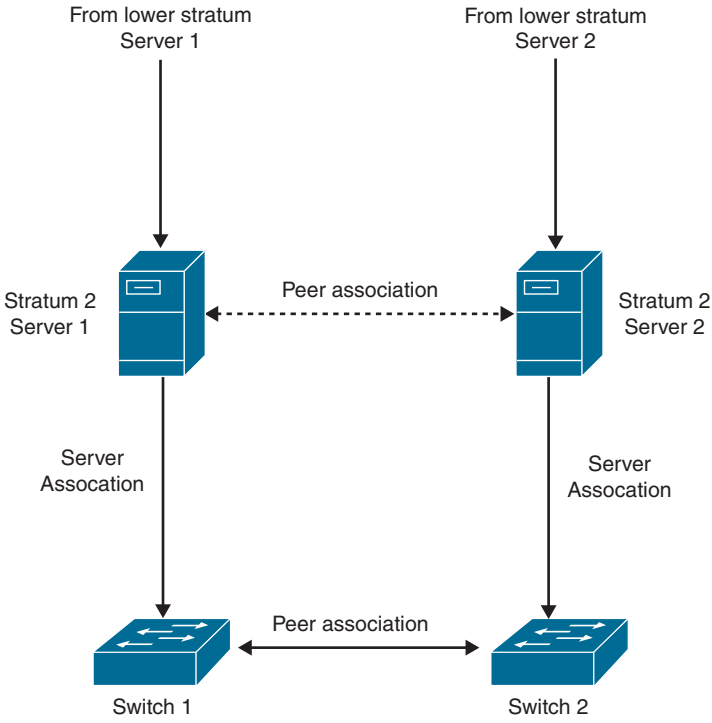



FIGURE 4.1 *NTP Peer and Server Associations*

Let us now look at the NTP configuration and verification steps:

1. Configure a client to use an NTP server by using the **ntp server** *ip-address* [**prefer**] [**source** *interface-id*] command.
2. When the device is acting as an NTP server, use the **ntp master** *stratum-number* command to set the stratum for a device statically. This command is only needed if the Cisco device is the official authoritative time source and is using its own clock as the trusted timepiece. Otherwise, it is not needed.
3. Verify the status of the NTP service by using the **show ntp status** command.
4. Verify the configuration by using the **show ntp associations** command.
5. For the NTP peer configuration, use the **ntp peer** *ip-address* command on both devices.

Figure 4.2 shows the NTP configuration used for the following example.

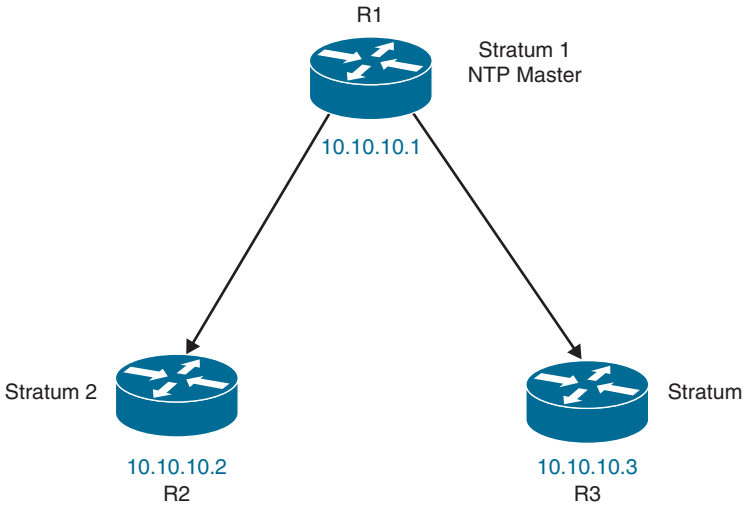



FIGURE 4.2 **NTP Server Configuration**

Example 4.1 shows how to configure three routers for NTP. R1 will be the NTP master and will use its own clock as the official timepiece, and R2 and R3 will be configured as NTP peers and will receive their time from R1.

In Example 4.1, there are two stratums: stratum 1 (R1) and stratum 2 (R2 and R3). You can assume for this example that R1 is directly attached to an authoritative time source. R2 and R3 are configured to query R1. If you had many more clients, you could have more stratums—up to stratum 15. To statically set a stratum as an NTP server, you use the command **ntp master** *stratum-number*. You use the command **ntp server** *ip-address* [**prefer**] [**source** *interface-id*] to configure the client. Setting the source interface is optional; if you set it, you specify the source IP address for queries for that server. You can set up multiple NTP servers for redundancy and use the **prefer** keyword to indicate the NTP server to use for time synchronization. You verify the status and stratum by using the **show ntp status** command. The output of this command shows information like the frequency and precision of the clock, NTP uptime, reference time, and polling interval and time since the last update.

EXAMPLE 4.1 **Configuring and Verifying NTP**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# ntp master ?
  <1-15>  Stratum number
  <cr>    <cr>
```

```
R1(config)# ntp master 1
R1(config)# end
R1#

R2#
R2# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R2(config)# ntp server 10.10.10.1 ?
  burst    Send a burst when peer is reachable (Default)
  iburst   Send a burst when peer is unreachable (Default)
  key      Configure peer authentication key
  maxpoll  Maximum poll interval
  minpoll  Minimum poll interval
  prefer   Prefer this peer when possible
  source   Interface for source address
  version  Configure NTP version
  <cr>     <cr>

R2(config)# ntp server 10.10.10.1
R2(config)# end
R2#

R3#
R3# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R3(config)# ntp server 10.10.10.1 prefer
R3(config)# end
R3#

R1# show ntp status
Clock is synchronized, stratum 1, reference is .LOCL.
nominal freq is 1000.0003 Hz, actual freq is 1000.0003 Hz, precision
is 2**16
ntp uptime is 128600 (1/100 of seconds), resolution is 1000
reference time is E4A4ECD4.C53226BD (12:23:00.770 UTC Fri Jul 23 2021)
clock offset is 0.0000 msec, root delay is 0.00 msec
root dispersion is 7937.55 msec, peer dispersion is 7937.50 msec
loopfilter state is 'CTRL' (Normal Controlled Loop), drift is
0.000000000 s/s
system poll interval is 16, last update was 3 sec ago.
R1#
R1#
R1# show ntp associations

  address       ref clock   st  when  poll reach  delay  offset  disp
*~127.127.1.1   .LOCL.       0     4    16   377  0.000   0.000   0.246
 * sys.peer, # selected, + candidate, - outlyer, x falseticker, ~
configured
```

```
R1#
 R2# !Similarly to R1, this shows us the status of the clock, the
stratum level, delays and other information
R2# show ntp status
Clock is synchronized, stratum 2, reference is 10.10.10.1
nominal freq is 1000.0003 Hz, actual freq is 1000.2067 Hz, precision
is 2**15
ntp uptime is 276100 (1/100 of seconds), resolution is 1000
reference time is E4A4FBC2.ABE8A834 (03:26:42.671 UTC Fri Jul 23 2021)
clock offset is 1.9363 msec, root delay is 2.82 msec
root dispersion is 29.19 msec, peer dispersion is 4.24 msec
loopfilter state is 'CTRL' (Normal Controlled Loop), drift is
-0.000206454 s/s
system poll interval is 128, last update was 425 sec ago.
R2#
R2# !This shows a streamlined version of the NTP server status and
delay. The address 10.10.10.1 shows the NTP association of R2.
R2# show ntp associations

  address      ref clock  st   when   poll reach  delay  offset   disp
*~10.10.10.1    .LOCL.       1     46    128   377  1.022   1.936  4.248
 * sys.peer, # selected, + candidate, - outlyer, x falseticker, ~
configured
R2#
R3# !As seen on R1 and R2, this shows us the status of the clock, the
stratum level, delays and other information
R3# show ntp status
Clock is synchronized, stratum 2, reference is 10.10.10.1
nominal freq is 1000.0003 Hz, actual freq is 1000.4177 Hz, precision
is 2**16
ntp uptime is 99200 (1/100 of seconds), resolution is 1000
reference time is E4A4F263.455DBDD7 (07:46:43.270 UTC Fri Jul 23 2021)
clock offset is 8.6126 msec, root delay is 3.48 msec
root dispersion is 7938.59 msec, peer dispersion is 437.56 msec
loopfilter state is 'CTRL' (Normal Controlled Loop), drift is
-0.000417468 s/s
system poll interval is 64, last update was 8 sec ago.
R3#
R3#! The address 10.10.10.1 shows the NTP association of R3 similarly
like we saw on R2
R3# show ntp associations
  address      ref clock   st   when   poll reach  delay  offset   disp
*~10.10.10.1   .LOCL.       1     58    128   377  1.106 762.748  4.998
 * sys.peer, # selected, + candidate, - outlyer, x falseticker,
~ configured
R3#
```

Figure 4.3 shows the NTP peer configuration and verification for the next example.

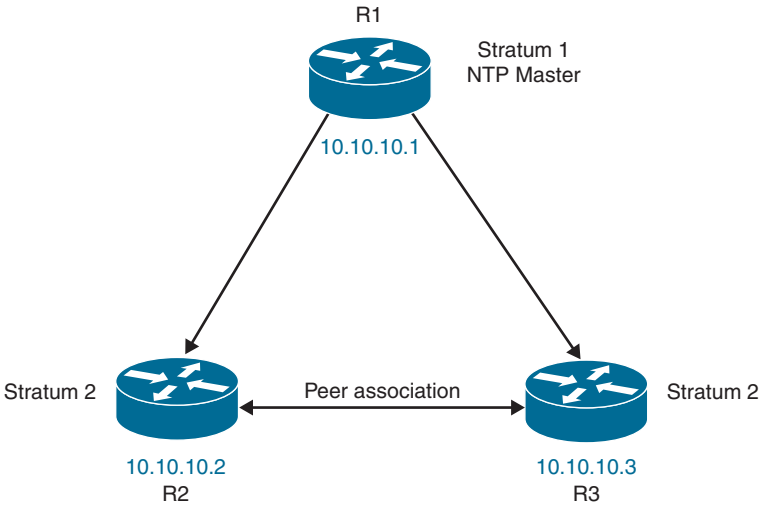



FIGURE 4.3 **NTP Peer Configuration**

Example 4.2 shows the configuration of an NTP peer relationship. In this example, R2 and R3 are peers with each other, and they query each other and move their time toward each other.

EXAMPLE 4.2 **Configuring NTP Peers**

```
R2#
R2# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R2(config)# ntp peer 10.10.12.3
R2(config)# end
R2#

R3#
R3# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R3(config)# ntp peer 10.10.12.2
R3(config)# end
R3#
```

**ExamAlert**

For the ENCOR exam, you should understand the need for NTP in a network and the two ways you can protect your NTP infrastructure: by using NTP authentication and by implementing NTP access lists.

NTP can be an easy target in a network, and you need to take a couple measures to secure your NTP infrastructure. The first involves setting up NTP authentication so the client can authenticate the server. It is important to note that the client authenticates the server; the server does not authenticate the client. Cisco devices support MD5 authentication for NTP.

The steps for configuring NTP authentication are as follows:

1. Define the NTP authentication key by using the **ntp authentication-key** *key-id* **md5** *key-string* command.
2. Enable authentication by using the **ntp authenticate** command.
3. Tell the device which key is valid for NTP authentication by using the **ntp trusted-key** *key-id* command.
4. Specify the NTP server that requires authentication by using the **ntp server** *server-ip-address* **key** *key-id* command. Note that the server is authenticated by the client.
5. Verify that the clock is still synchronized by using the **show ntp status** command.

Example 4.3 demonstrates NTP authentication. The NTP authentication key is defined as *key 1* with the key string *ExamCram123*. This example points to NTP server 100.1.1.1 using *key 1*.

EXAMPLE 4.3 **Authenticating NTP**

```
R2#
R2# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R2(config)# ntp authentication-key 1 md5 ExamCram123
R2(config)# ntp authenticate
R2(config)# ntp trusted-key 1
R2(config)# ntp server 100.1.1.1 key 1
R2(config)# end
R2#
```

The second way of securing NTP infrastructure involves using an access list on devices that synchronize their time with an external server. Because NTP does not authenticate clients, the devices continue to respond to authenticated requests, and it is therefore important to use an access list to limit NTP access.

A server uses access control lists (ACLs) to control who the server will give time to, based on the SRC IP address of the NTP messages from the client.

If the SRC IP address matches the ACL on the server, the server responds with the time. If it does not match, the server drops the message.

There are four restrictions that can be used with the **ntp access-group** global configuration command:

- **peer:** This restriction allows time synchronization requests and control queries. The device is allowed to synchronize itself to the servers specified in the access list.
- **serve:** This restriction allows time synchronization requests and control queries. The device is not allowed to synchronize itself to servers specified in the access list.
- **serve-only:** With this restriction, only synchronization requests are allowed.
- **query-only:** With this restriction, only control queries are allowed.

It is important to set a specific interface to serve as the source interface for NTP. The source of NTP packets is the same as the interface that the packet was sent out on. When implementing NTP authentication and an NTP access list, it is wise to set the source interface. Ideally, you should set a loopback interface as the NTP source as it will never be down, whereas a physical interface may go down.

Example 4.4 shows how to set Lookback 0 as the NTP source interface for predictability. It specifies that the device must allow **peer** requests and specifies server 100.1.1.1 in access list 10, from which it accepts responses.

EXAMPLE 4.4 **Creating an NTP Access List**

```
R2#
R2# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R2(config)# ntp source Loopback 0
R2(config)# ntp server 100.1.1.1
R2(config)# access-list 10 permit 100.1.1.1
R2(config)# ntp access-group peer 10
R2(config)# end
R2#
```

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. True or false: Accurate time on a network device is important in troubleshooting scenarios for the analysis of system logs.
   - ❍ **A.** True
   - ❍ **B.** False

2. Which of the following does NTP use to determine the number of hops to the authoritative time source?
   - ❍ **A.** Hop count
   - ❍ **B.** Administrative distance
   - ❍ **C.** Layer
   - ❍ **D.** Stratum

3. Which of the following commands is used to enable a Cisco device as an authoritative NTP server?
   - ❍ **A. ntp server**
   - ❍ **B. ntp master**
   - ❍ **C. ntp peer**
   - ❍ **D. ntp association**

4. True or false: When setting up NTP authentication, the NTP client authenticates the NTP server to prevent unauthorized NTP communication.
   - ❍ **A.** True
   - ❍ **B.** False

## Answers

1. **A** is correct. Accurate time assists in troubleshooting issues with network devices by identifying time-specific events in system logs.
2. **D** is correct. NTP uses a stratum to determine how many NTP hops away a network device is from the authoritative time source.
3. **B** is correct. To configure a Cisco device as an authoritative NTP server, you use the **ntp master** *stratum-number* command.
4. **A** is correct. The NTP client authenticates the NTP server to prevent authorized hosts from communicating with the NTP service on the device.

# Network Address Translation (NAT)

This section looks at the configuration of Network Address Translation (NAT) for IP address conservation and enhanced security. It describes the benefits of configuring NAT in a network environment as well as the various types of NAT.

NAT enables private IP networks that use nonregistered IP addresses to connect to the Internet. NAT operates on devices, such as routers and firewalls, that typically connect networks. Before packets are forwarded to another network, NAT translates the private (that is, non-globally unique) addresses in the internal network into a globally routable address space. NAT can provide better security by effectively hiding the internal network behind a single globally routable address. It does this by advertising to the outside world only that one address for the entire network. NAT also provides a graceful renumbering strategy for organizations that are changing service providers or renumbering into classless interdomain routing (CIDR) blocks. Finally, for simple load sharing of traffic, you can map a single global IP address with many local IP addresses by using TCP load distribution.

**ExamAlert**

For the ENCOR exam, you need to know the three types of NAT (static, dynamic, and overloading) as well as the terms associated with NAT configuration and verification.

At a high level, NAT can be classified into three types:

- **Static NAT:** Static NAT facilitates one-to-one mapping between local and global addresses. That is, one public address is required for every private address.
- **Dynamic NAT:** Dynamic NAT facilitates one-to-one mapping between local and global addresses from a pool of registered IP addresses. One public address is required for every private address.
- **Overloading (PAT):** Overloading makes it possible to map multiple unregistered IP addresses to a single registered IP address (many to one) by using different ports. It is also known as Port Address Translation (PAT). PAT is the best solution for connecting enterprise network users to the Internet.

To better understand NAT configuration and verification, you need to have a good understanding of the following terms:

- **Inside local address:** An IP address that is assigned to a host on the inside network. For the ENCOR exam, the inside local address is a private IP address, as specified in RFC 1918.
- **Inside global address:** A legitimate IP address representing one or more inside local IP addresses to the outside world. For the ENCOR exam, the inside global address is a public IP address on the Internet that has been assigned to the company that is performing NAT.
- **Outside local address:** The IP address of an outside host as it appears to the inside network. For the ENCOR exam, the outside local address is the same IP address as the outside global address.
- **Outside global address:** The IP address assigned to a host on the outside network by the owner of the host. For the ENCOR exam, the outside global address is a public IP address on the Internet of the device your hosts (users) are trying to reach.

Figure 4.4 shows the reference topology that is used in this section. It shows the inside and outside networks along with the router (R1) where you will make NAT configurations.

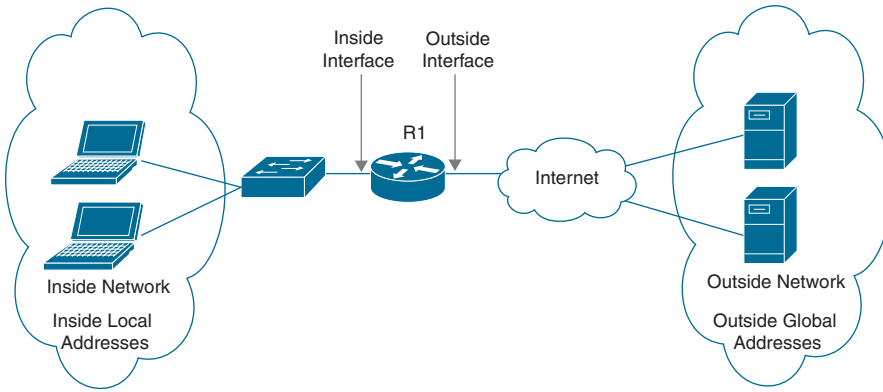



FIGURE 4.4 **NAT Configuration Topology**

The following sections provide examples of the different NAT configuration options.

# Static NAT

As mentioned earlier in this chapter, static NAT involves translating a global IP address to a local IP address, based on a static mapping between the global IP address and the local IP address. There are two types of static NAT: inside static NAT and outside static NAT.

With inside static NAT, you map an inside local (private RFC 1918) address to a single inside global (public) address. In this case, the private address is hidden from the outside world. These are the configuration steps:

1. Configure the outside interface by using the **ip nat outside** command.
2. Configure the inside interface by using the **ip nat inside** command.
3. Configure the inside static NAT by using the **ip nat inside source static** *inside-local-ip inside-global-ip* command.
4. Verify the configuration by using the **show ip nat translations** command.

Figure 4.5 shows the topology for a static NAT configuration.

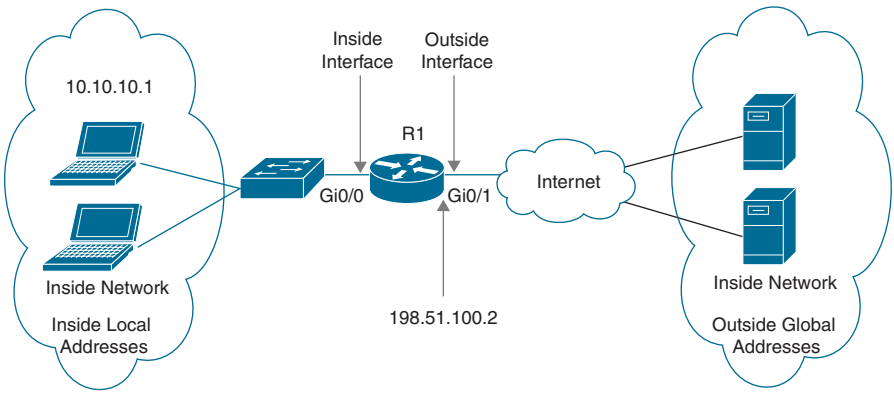



FIGURE 4.5 **Static NAT Topology**

Example 4.5 shows the configuration of static NAT. It translates one private IP address to a single public IP address. In this case, the private IP address 10.10.10.1 is mapped to public IP address 198.51.100.2. Private NAT is not very common as it requires one public IP address for each private IP address. Example 4.5 defines the inside and outside interfaces and then uses the **ip nat inside static** command to map the private IP address 10.10.10.1 to 198.51.100.2. The example shows how to verify the NAT mapping in the NAT translation table by using the command **show ip nat translations**.

EXAMPLE 4.5 **Configuring Static NAT**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# interface GigabitEthernet0/0
R1(config-if)# ip nat inside
R1(config-if)# interface GigabitEthernet0/1
R1(config-if)# ip nat outside
R1(config-if)# exit
R1(config)# ip nat inside source static 10.10.10.1 192.51.100.2
R1(config)# end
R1#
R1# show ip nat translations
Pro Inside global   Inside local     Outside local      Outside global
--- 192.51.100.2    10.10.10.1       ---                ---
R1#
```

With outside static NAT, you map an outside global (public) address to an outside local (private) address. In this case, the real external IP addresses are hidden from the inside hosts.

# Dynamic NAT

One of the major drawbacks of using static NAT is the number of global IP addresses needed to match the number of local IP addresses; static NAT is therefore not scalable in larger environments. In addition, static NAT is tedious and resource intensive on a device that has a lot of configuration entries.

Dynamic NAT provides address translation on an as-needed basis. The global IP address is from a pool, and the dynamic translation stays in the translation table as long as traffic flows from the local address to the global address. The translation is removed once translation stops and the timeout period expires. With dynamic NAT, you still need one public IP address for every private IP address. Dynamic NAT can operate as either inside NAT or outside NAT.

These are the configuration steps for dynamic NAT:

1. Configure the outside interface by using the **ip nat outside** command.
2. Configure the inside interface by using the **ip nat inside** command.
3. Specify the object to be translated by using a standard or extended ACL referenced by either number or name. A standard ACL states the source IP addresses that could be translated, and an extended ACL allows for conditional translation based on protocol, port, or source or destination IP addresses. The ACL is used to define the inside local addresses.

4. Define the global pool of IP addresses (that is, inside global addresses) by using the **ip nat pool** *nat-pool-name starting-ip ending-ip* **prefix-length** *prefix-length* command.
5. Configure the inside hosts' address translation to the NAT pool by using the command **ip nat inside source list** *acl* **pool** *nat-pool-name*.
6. Verify the configuration by using the **show ip nat translations** command.
7. If you would like to change the timeout for NAT translations from the default of 24 hours, use the **ip nat translation timeout** *seconds* command.
8. Clear the NAT translation table by using the command **clear ip nat translation** {*ip-address* | *}. This could interrupt traffic flow on active connections as a new global IP address may be assigned from the pool.

# Port Address Translation (PAT)

You can conserve IP addresses in the inside global address pool by allowing a device to use one global address for many local addresses. This type of NAT configuration is referred to as *overloading*. When overloading is configured, the device maintains information from higher-level protocols, such as TCP or UDP port numbers. This action translates the global address back to the correct local address. When multiple local addresses map to one global address in a translation, the TCP or UDP port numbers of each inside host distinguish between local addresses.

PAT addresses the limitations of dynamic NAT by working well with a large number of global IP addresses and ensuring that the number of global IP addresses is sufficient for the local IP addresses that you are translating.

You configure PAT as follows:

1. Configure the outside interface by using the **ip nat outside** command.
2. Configure the inside interface by using the **ip nat inside** command.
3. Specify the traffic to be translated by using a standard or extended ACL referenced by number or name. This ACL is used to define the inside local addresses.
4. Configure PAT by using the **ip nat inside source list** *acl* {**interface** *interface-id* | **pool** *nat-pool-name*} **overload** command. Specifying an interface involves using the primary IP address assigned to that interface. Specifying a NAT pool requires creating the NAT pool that uses those IP addresses as global addresses.

5. Verify the configuration by using the **show ip nat translations** command.

Example 4.6 shows the configuration of PAT. It overloads a single IP address that is on GigabitEthernet0/1. The public IP address that is on this interface, 198.51.100.2, is used.

EXAMPLE 4.6 **Configuring PAT**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# ip access-list standard PAT-ACL
R1(config-std-nacl)# permit 10.10.10.0 0.0.0.255
R1(config-std-nacl)# exit
R1(config)# interface GigabitEthernet 0/0
R1(config-if)# ip nat inside
R1(config-if)# interface GigabitEthernet 0/1
R1(config-if)# ip nat outside
R1(config-if)# exit
R1(config)# ip nat inside source list PAT-ACL interface GigabitEther-
net 0/1 overload
R1(config)# end
R1#
R1#
R1# show run int GigabitEthernet 0/0
Building configuration...

Current configuration : 177 bytes
!
interface GigabitEthernet0/0
 ip address 10.10.10.1 255.255.255.0
 ip nat inside
 ***Output Omitted***
end

R1# show run int GigabitEthernet 0/1
Building configuration...

Current configuration : 175 bytes
!
interface GigabitEthernet0/1
 ip address 198.51.100.2 255.255.255.0
 ip nat outside
 ***Output Omitted***
end
R1#
```

When using PAT with a single public IP address, you are limited to 64,512 TCP source ports and 64,512 UDP ports. TCP and UDP each support 65,536 ports per IP address (public IP address for our purposes), but the first 1,024 well-known (privileged) ports are not used. As a result, overloading on a NAT pool (a pool of public IP address) is common in enterprise deployments as each public IP address can support 64,512 translations.

Example 4.7 shows the configuration of PAT with a pool of public IP addresses. This example is very similar to Example 4.6. The main difference is that it defines the global pool of IP addresses using the command **ip nat pool** *nat-pool-name starting-ip ending-ip* **prefix-length** *prefix-length*. In addition, when you configure NAT overload, you overload on the public pool and not the outside interface. This example defines the inside local addresses in an ACL called INSIDE-ADDRESSES, and the public pool is called PUBLIC-ADDRESSES-POOL. Notice in the **show ip nat translation** output in this example that the port for the inside global entries is unique. Only public IP address 198.51.100.5 is being used with different ports. As these ports are exhausted, different traffic flow will use another public IP address from the PUBLIC-ADDRESSES-POOL pool.

EXAMPLE 4.7 **Configuring PAT with a Dynamic Pool**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# ip access-list standard INSIDE-ADDRESSES
R1(config-std-nacl)# permit 10.10.10.0 0.0.0.255
R1(config-std-nacl)# exit
R1(config)# interface GigabitEthernet 0/0
R1(config-if)# ip nat inside
R1(config-if)# interface GigabitEthernet 0/1
R1(config-if)# ip nat outside
R1(config-if)# exit
R1(config)# ip nat pool PUBLIC-ADDRESSES-POOL 198.51.100.20
198.51.100.29 prefix-length 24
R1(config)# ip nat inside source list INSIDE-ADDRESSES pool PUBLIC-
ADDRESSES-POOL overload
R1(config)# end
```

```
R1#
R1# show ip nat translation
Pro Inside global      Inside local       Outside local  Outside global
icmp 198.51.100.21:22  10.10.10.1:22      10.10.20.1:22  10.10.20.1:22
tcp 198.51.100.21:51697 10.10.10.1:51697 10.10.20.1:23  10.10.20.1:23
icmp 198.51.100.21:6   10.10.10.2:6       10.10.20.1:6   10.10.20.1:6
tcp 198.51.100.21:23678 10.10.10.2:23678 10.10.20.1:23  10.10.20.1:23
```

A NAT virtual interface (NVI), which removes the condition to configure an interface as either a NAT inside or a NAT outside interface, was introduced in IOS 12.3(14)T. An interface can be configured to use or not use NAT. You can think of traditional NAT as domain-based NAT, where, as you specify an inside or outside interface, you are also specifying an inside domain and an outside domain. When you specify a domain, translation rules are applied before or after route decisions are applied—depending on whether you are translating from inside to outside or outside to inside. However, with an NVI, translation rules are applied to a domain only after a route decision for an NVI is applied.

An NVI provides simplification in the sense that when a NAT pool is shared for translating packets from multiple networks connected to a NAT device, an NVI is created, and a static route is configured that forwards all packets addressed to the NAT pool to the NVI. Standard interfaces connected to the various networks are configured to determine if the traffic, originating from and received on the interfaces, needs to be NAT translated.

At a high level, an NVI basically provides two benefits:

- A NAT table is maintained per interface, which provides better performance and scalability.
- The need for domain-specific NAT configuration is eliminated.

To configure an NVI, you use the interface configuration command **ip nat enable** on the inside and outside interfaces where NAT needs to be performed. You basically eliminate the need to configure the **inside** and **outside** keywords on interfaces. All the other commands are similar to those used in configuring traditional NAT.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. True or false: Rather than create a new IP host address on an inside local client, you can use NAT to assist in renumbering a network.
   - ❍ **A.** True
   - ❍ **B.** False

2. Which of the following terms refers to the public IP address assigned to a host on the outside network by the owner of the host?
   - ❍ **A.** Inside local address
   - ❍ **B.** Inside global address
   - ❍ **C.** Outside local address
   - ❍ **D.** Outside global address

3. Which type of NAT allows for one-to-one mapping of unregistered IP addresses to registered IP addresses from a pool of registered IP addresses?
   - ❍ **A.** Inside static NAT
   - ❍ **B.** Outside static NAT
   - ❍ **C.** Dynamic NAT
   - ❍ **D.** PAT

## Answers

1. **A** is correct. NAT also provides a graceful renumbering strategy for organizations that are changing service providers or numbering into classless interdomain routing (CIDR) blocks.
2. **D** is correct. An outside global address is a public IP address assigned to a host on the outside network by the owner of the host.
3. **C** is correct. Dynamic NAT is used for mapping unregistered IP addresses to registered IP addresses from a pool of registered IP addresses.

# First-Hop Redundancy Protocols (FHRPs)

In this section, we take a look at using FHRPs such as Hot Standby Router Protocol (HSRP), Virtual Router Redundancy Protocol (VRRP), and Gateway Load Balancing Protocol (GLBP) to provide a redundant gateway for hosts. It covers the configuration and verification of these protocols but does not go into much detail on the basics of these technologies and the problems that they solve. The basics and designs of these FHRPs are covered in Chapter 19, "Enterprise Network Design Principles." If you need a refresher on these protocols before getting into the configuration, it may be worthwhile to skip ahead and read that chapter before coming back to look at the configuration and verification of HSRP, VRRP, and GLBP here.

Hot Standby Router Protocol (HSRP) is one of the most commonly used gateway redundancy protocols. It is a Cisco-proprietary protocol that can be used in a group of routers or multilayer switches for selecting an active device and a standby device. In a group of device interfaces, the active device is used for routing packets, and the standby device takes over when the active device fails or when preset conditions are met.

> **ExamAlert**
>
> For the ENCOR exam, you need to have a clear understanding of how HSRP is configured, the optional configuration options, and the advantages it provides to end hosts.

Once you enable HSRP on routers or multilayer switch interfaces by specifying the virtual IP address for a group, the virtual IP address and an associated virtual MAC address are available for end hosts to use. End hosts point to the virtual IP address, which serves as the default gateway, and the host then learns the virtual MAC address via Address Resolution Protocol (ARP).

A number of mandatory and optional steps are involved in an HSRP configuration. Only step 2 in the following list, where you define the HSRP group with the **standby** command, is mandatory. The other steps enhance the configuration or make an HSRP deployment more resilient. Let us look at these steps:

1. Configure the HSRP version by using the **standby version {1 | 2}** command.
2. Define the HSRP group by using the **standby** [*group-number*] **ip** [*virtual ip address*] command.

3. Set the HSRP priority by using the **standby** [*group-number*] **priority** [*priority*] command.
4. To enable object tracking to decrement the priority when the state of an object is false, use the command **standby** [*group-id*] **track** [*object-id*] **decrement** *decrement-value*. The decrement value needs to be high enough so that when the router is removed from the group, the value is lower than that of the other HSRP router in the group.
5. Configure preemption by using the **standby** [*group-number*] **preempt** [**delay** [**minimum** *seconds*] [**reload** *seconds*] [**sync** *seconds*]] command.
6. Configure HSRP timers by using the **standby** *[group-number]* **timers** {*seconds* | **msec** *milliseconds*} command.
7. Conduct HSRP authentication by using the command **standby** *[group-number]* **authentication** [**text** *string* | **md5** {**key-chain** *key-chain* | **key-string** *key-string*}].
8. Verify the configuration by using the **show standby** [*interface-id*] **brief** command.

The active router in the group is responsible for the virtual address. The other router in the group is in the standby state and monitors the active router. Additional routers are in the HSRP listen state. If the active router fails, the standby router assumes the active state. If the standby routers fail to become the active router, then another router is elected as the standby router.

HSRP devices are always in one of the following states:

- **Active:** The device is performing packet transfer functions.
- **Init or disabled:** The device is not yet ready or able to participate in HSRP.
- **Learn:** The device has not determined the virtual IP address and has not yet seen an authenticated hello message from the active device.
- **Listen:** The device is receiving hello messages.
- **Speak:** The device is sending and receiving hello messages.
- **Standby:** The device is prepared to accept the packet transfer function if the active device fails.

In the following configuration example, you will encounter the following configuration options:

- **HSRP priority:** This option determines which devices are the active devices in the group. The priority is determined first by the configured value and then by the IP address. In the event of a tie, the primary IP addresses are compared, and the device with the higher IP address has the priority. In each case, the higher value is of greater priority.
- **HSRP preemption:** This option enables the HSRP router with the highest priority to immediately become the active router once it is available. When preemption is enabled, the switch with the highest priority becomes the new active device. Preemption can be enabled with the **standby preempt** command.

Figure 4.6 shows the reference topology for the following HSRP configuration example.

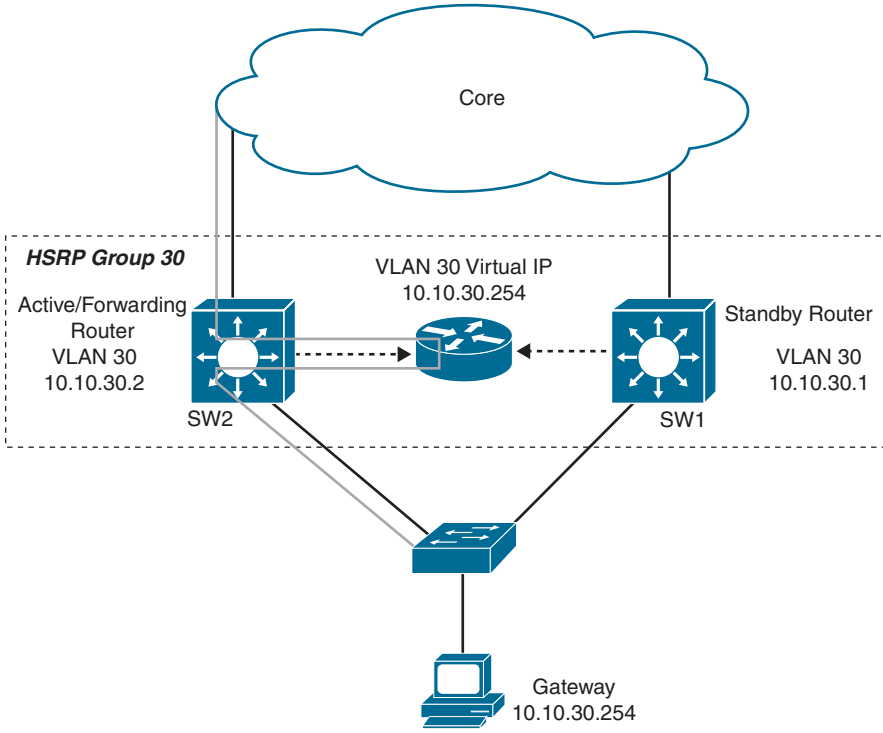



FIGURE 4.6 **HSRP Configuration Topology**

Example 4.8 shows the basic configuration and verification of HSRP between two Layer 3 switches for VLAN 30. It configures HSRP group 30 with the virtual IP address 10.10.30.254, which serves as the gateway for the host. This example also configures preemption to allow a more preferred router to take the active router status from an inferior active router if needed. Preemption is an optional configuration. For verification, the **show standby** output shows the virtual IP address and the active router. (The virtual IP address and the active router both have the default priority 100, so the highest IP address—in this case, 10.10.30.2—wins.) You can see the preemption status and the active virtual MAC address. The **show standby brief** output shows a summary view of the active and standby routers, the priority, and the virtual IP address.

**ExamAlert**

For the ENCOR exam, you should know how the active virtual MAC address is determined. For example, for HSRPv1, the virtual MAC address is 0000:0c07:acXX (with XX representing the HSRP group, in hexadecimal). In this case, the 1e in the MAC address 0000:0c07:ac1e represents 30 (for HSRP group number 30). If you were using HSRPv2, the virtual MAC address would be 0000:0c9f:fXXX (with XXX representing the HSRP group, in hexadecimal).

EXAMPLE 4.8 **Configuring and Verifying HSRP**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# interface vlan 30
SW1(config-if)# ip address 10.10.30.1 255.255.255.0
SW1(config-if)# standby 30 ip 10.10.30.254
SW1(config-if)# standby 30 preempt
SW1(config-if)# end
SW1#

SW2#
SW2# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW2(config)# interface vlan 30
SW2(config-if)# ip address 10.10.30.2 255.255.255.0
SW2(config-if)# standby 30 ip 10.10.30.254
SW2(config-if)# standby 30 preempt
SW2(config-if)# end
SW2#
```

```
SW2#
SW2# show standby
Vlan30 - Group 30
  State is Active
    2 state changes, last state change 00:02:55
  Virtual IP address is 10.10.30.254
  Active virtual MAC address is 0000.0c07.ac1e (MAC In Use)
    Local virtual MAC address is 0000.0c07.ac1e (v1 default)
  Hello time 3 sec, hold time 10 sec
    Next hello sent in 0.368 secs
  Preemption enabled
  Active router is local
  Standby router is 10.10.30.1, priority 100 (expires in 10.592 sec)
  Priority 100 (default 100)
  Group name is "hsrp-Vl30-30" (default)
SW2#

SW2# show standby brief
                     P indicates configured to preempt.
                     |
Interface   Grp  Pri P State   Active    Standby         Virtual IP
Vl30        30   100 P Active  local     10.10.30.1      10.10.30.254
SW2#
```

# Virtual Router Redundancy Protocol (VRRP)

VRRP is a standards-based FHRP that serves as an alternative to the Cisco-proprietary HSRP. It is similar to HSRP in terms of operation and configuration. The VRRP master role is analogous to the HSRP active role, and the VRRP backup role is analogous to the HSRP standby role. A VRRP group has one master device and one or more backup devices. The device with the highest priority becomes the master, and the priority can range from 0 to 255.

Unlike HSRP, VRRP allows you to use an IP address of one of the physical VRRP group members as the virtual IP address. The device with that physical address is the VRRP master whenever it is available. The default advertisement interval is 1 second, with a 3-second hold time; in comparison, HSRP has a default 3-second hello time and 10-second hold time.

A number of mandatory and optional steps are involved in a VRRP configuration. In the following steps, only step 1, where you define the VRRP group

with the **vrrp** command, is mandatory. The other steps enhance the configuration or make a VRRP deployment more resilient. These are the steps:

1. Define the VRRP group by using the **vrrp** [*group-id*] **ip** [*virtual ip address*] command.
2. Define the VRRP priority by using the command **vrrp** [*group-id]* **priority** *priority*].
3. To enable object tracking to decrement the priority when the state of an object is false, use the command **vrrp** [*group-id*] **track** [*object-id*] **decrement** [*decrement-value*]. This decrement value needs to be high enough so that when it is removed from priority, the value is lower than that of the other VRRP router in the group.
4. Authenticate VRRP by using the command **vrrp** [*instance-id*] **authentication** {**text** string | **md5** {**key-chain** *key-chain* | **key-string** *key-string*}}.
5. Finally, verify the configuration by using the **show vrrp** [**brief**] command.

Figure 4.7 shows the reference VRRP topology for the following VRRP configuration example.

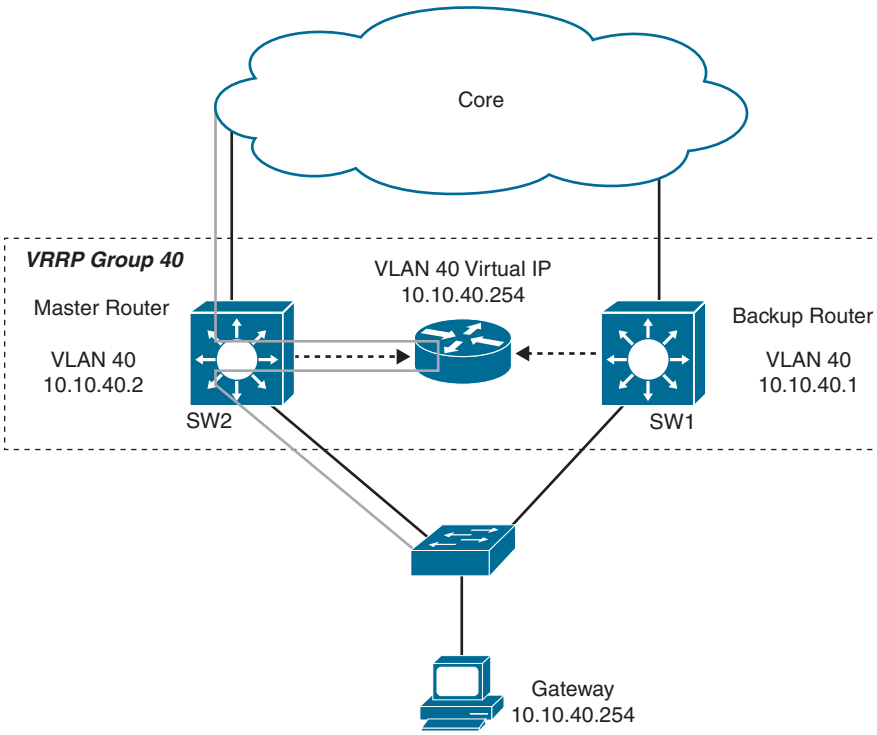



FIGURE 4.7 **VRRP Configuration Topology**

Example 4.9 shows the basic configuration and verification of VRRP between two Layer 3 switches for VLAN 40. It shows the configuration of VRRP group 40 with virtual IP address 10.10.40.254. This serves as the gateway for the host. VRRP enables preemption by default to allow a more preferred router to take the master router status from an inferior master router, if needed. For verification, the **show vrrp** output shows the virtual IP address and the master router. (The virtual IP address and the master router both have the default priority 100, so the highest IP address—in this case, 10.10.40.2—wins.) Priority can be between 1 and 254. The **show vrrp brief** output shows a summary view of the master and virtual routers, the priority, and the virtual IP address.

**ExamAlert**

For the ENCOR exam, you need to know how the VRRP virtual MAC address is determined. For example, say that the virtual MAC address is 0000:5e00:01XX (with XX representing the VRRP group, in hexadecimal). In this case, the 28 in the MAC address 0000:5e00:0128 represents 40 (for VRRP group number 40).

EXAMPLE 4.9 **Configuring and Verifying VRRP**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# interface vlan40
SW1(config-if)# ip address 10.10.40.1 255.255.255.0
SW1(config-if)# vrrp 40 ip 10.10.40.254
SW1(config-if)# end
SW1#

SW2#
SW2# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW2(config)# interface vlan 40
SW2(config-if)# ip address 10.10.40.2 255.255.255.0
SW2(config-if)# vrrp 40 ip 10.10.40.254
SW2(config-if)# end
SW2#

SW2# show vrrp
Vlan40 - Group 40
  State is Master
  Virtual IP address is 10.10.40.254
  Virtual MAC address is 0000.5e00.0128
  Advertisement interval is 1.000 sec
  Preemption enabled
```

```
  Priority is 100
  Master Router is 10.10.40.2 (local), priority is 100
  Master Advertisement interval is 1.000 sec
  Master Down interval is 3.609 sec

SW2# sho vrrp brief
Interface   Grp Pri Time  Own Pre State   Master addr   Group addr
Vl40        40  100 3609        Y  Master  10.10.40.2    10.10.40.254
SW2#
```

# Gateway Load Balancing Protocol (GLBP)

**ExamAlert**

For the ENCOR exam, make sure you have a complete understanding of GLBP, including how it is similar to HSRP and VRRP and how it differs in operation.

GLBP is a Cisco-proprietary protocol. Like HSRP and VRRP, it provides redundant default gateways; however, it allows you to use all the default gateways at the same time instead of just using one, which is the limit for HSRP and VRRP. With GLBP, you can maximize the use of your links and gateways by load balancing traffic across different links and gateways within the same subnet/virtual local area network (VLAN). This is accomplished by using one virtual IP address and many MAC addresses. The gateways share the virtual IP address, but each gateway uses a different MAC address. So, an end user who uses ARP with the MAC address of the gateway is provided with one of the many virtual MAC addresses so that traffic can be forwarded to that specific gateway. The default is to return the MAC addresses in a round-robin fashion, but you can change it to other options if you wish (although they are beyond the scope of the ENCOR exam and this chapter).

A number of mandatory and optional steps are involved in a GLBP configuration. Only step 1, where you define the GLBP instance with the **glbp** command, is mandatory. The other steps enhance the configuration or make a GLBP deployment more resilient. These are the steps:

1. Define the GLBP instance by using the command **glbp** [*group-id*] **ip** [virtual ip *address*].
2. Configure GLBP preemption to allow for a more preferred router to take the active virtual gateway status from an inferior active GLBP router. To do so, use the command **glbp** [*instance-id*] **preempt**.

3. Define the GLBP priority by using the command **glbp** [*group-id*] **priority** [*priority*].
4. Define the GLBP timers by using the command **glbp** [*group-id*] **timers** {*hello-seconds* | **msec** *hello-milliseconds*} {*hold-seconds* | **msec** *hold-milliseconds*}.
5. Authenticate GLBP by using the command **glbp** [*group-id*] **authentication** {**text** *string* | **md5** {**key-chain** *key-chain* | **key-string** *key-string*}}.
6. If desired, change the load-balancing method by using the command **glbp** [*group-id*] **load-balancing** {**host-dependent** | **round-robin** | **weighted**}.
7. Verify the configuration by using the **show glbp** [**brief**] command.

Figure 4.8 shows the reference GLBP topology for the following GLBP configuration example.

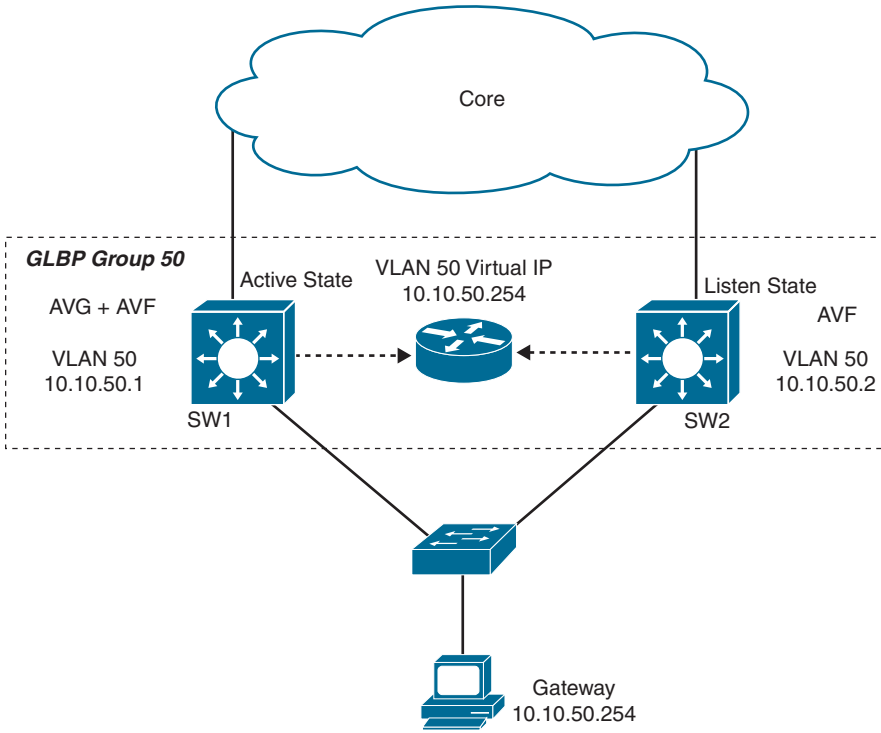



FIGURE 4.8 **GLBP Configuration Topology**

Example 4.10 shows the basic configuration and verification of GLBP between two Layer 3 switches for VLAN 50. It configures GLBP group 50 with the

virtual IP address 10.10.50.254. This serves as the gateway for the host. The switches in Group 50 elect SW2 as the active virtual gateway (AVG). There is one AVG per group, and the other switches are active virtual forwarders (AVFs). The AVG assigns a virtual MAC address to any other switch in group 50. All switches in a group are AVF, including, in this case, SW1, which is also the AVG. ARP requests from the host will be responded to by the AVG with the virtual MAC address of one of the AVFs. Therefore, all switches in the group will be used.

Preemption with GLBP is disabled by default. For verification, the output of the **show glbp** command shows the virtual IP address and the AVG router (with the default priority of 100). If the AVG fails, the AVF takes over; if there are multiple AVFs with the same priority, the one with the highest IP address wins and becomes the new AVG. Priority can be between 1 and 255. The output of the **show glbp brief** command shows a summary of the active and standby routers, the priority, and the virtual IP address.

**ExamAlert**

For the ENCOR exam, you need to know how the GLBP virtual MAC address is determined. For example, say that the virtual MAC address is 0007:b400:XXYY (where XX represents the GLBP group number, in hexadecimal, and YY represents the AVF number). In this case, the 32 in the MAC address 0007:b400:3201 represents 50 (for GLBP group 50), and it is the same for the MAC address 0007:b400:3202. Only the 01 and 02 differentiate these MAC addresses; they represent the two AVFs.

EXAMPLE 4.10 **Configuring and Verifying GLBP**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# interface vlan 50
SW1(config-if)# ip address 10.10.50.1 255.255.255.0
SW1(config-if)# glbp 50 ip 10.10.50.254
SW1(config-if)# glbp 50 preempt
SW1(config-if)# end
SW1#

SW2#
SW2# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW2(config)# interface vlan 50
SW2(config-if)# ip address 10.10.50.2 255.255.255.0
```

```
SW2(config-if)# glbp 50 ip 10.10.50.254
SW2(config-if)# glbp 50 preempt
SW2(config-if)# end
SW2#

SW1# show glbp
Vlan50 - Group 50
  State is Active
    1 state change, last state change 00:14:14
  Virtual IP address is 10.10.50.254
  Hello time 3 sec, hold time 10 sec
    Next hello sent in 0.704 secs
  Redirect time 600 sec, forwarder timeout 14400 sec
  Preemption enabled, min delay 0 sec
  Active is local
  Standby is 10.10.50.2, priority 100 (expires in 9.280 sec)
  Priority 100 (default)
  Weighting 100 (default 100), thresholds: lower 1, upper 100
  Load balancing: round-robin
  Group members:
    5254.0005.8032 (10.10.50.1) local
    5254.0013.8032 (10.10.50.2)
  There are 2 forwarders (1 active)
  Forwarder 1
    State is Active
      1 state change, last state change 00:14:03
    MAC address is 0007.b400.3201 (default)
    Owner ID is 5254.0005.8032
    Redirection enabled
    Preemption enabled, min delay 30 sec
    Active is local, weighting 100
  Forwarder 2
    State is Listen
    MAC address is 0007.b400.3202 (learnt)
    Owner ID is 5254.0013.8032
    Redirection enabled, 599.296 sec remaining (maximum 600 sec)
    Time to live: 14399.296 sec (maximum 14400 sec)
    Preemption enabled, min delay 30 sec
    Active is 10.10.50.2 (primary), weighting 100 (expires in
11.072 sec)
SW1#
SW1#
SW1# show glbp brief
Interface Grp  Fwd Pri State  Address        Active router   Standby
                                                            router
Vl50      50   -   100 Active 10.10.50.254   local           10.10.50.2
Vl50      50   1   -   Active 0007.b400.3201 local           -
Vl50      50   2   -   Listen 0007.b400.3202 10.10.50.2      -
SW1#
```

# Object Tracking with FHRPs

An FHRP can be used with object tracking. Object tracking provides the flexibility to link an FHRP with other router components, such as static routes. For example, HSRP can be used with object tracking to take a particular action when an object state changes in the network. Object tracking can be taken a step further by tracking the output of an IP service level agreement (SLA) object and providing the instruction to an FHRP for action. IP SLAs are covered in Chapter 31, "IP SLA and DNA Center."

The object tracking feature allows you to create a tracked object that various clients can use to influence the client behavior when a tracked object changes. It enables you to track specific objects on a device—such as an interface line protocol state, IP routing, and route reachability—and then take action when the tracked object's state changes. The object tracking feature essentially allows you to increase the network's availability and shorten recovery time if an object state goes down.

Several clients can register their interest with the tracking process, can track the same object, and can take different actions when the object state changes. Clients include the following:

- Embedded Event Manager (EEM)
- HSRP
- VRRP
- GLBP

The following steps are involved in the configuration and verification of the tracking of routes in the routing table and the tracking on an interface line protocol state:

1. Track routes in the routing table, using the command **track** *object-number* **ip route** *route/prefix-length* **reachability**.
2. Track an interface line protocol's state by using the command **track** *object-number* **interface** *interface-id* **line-protocol**.
3. Verify the status of object tracking by using the command **show track** [*object-number*].

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which of the following first-hop redundancy protocols is considered standards based?
   - ❍ **A.** HSRP
   - ❍ **B.** VRRP
   - ❍ **C.** GLBP
   - ❍ **D.** None of the above

2. With HSRP, the active default gateway that serves client requests is provided from which of the following?
   - ❍ **A.** The active member's IP address
   - ❍ **B.** The standby members IP address
   - ❍ **C.** The virtual IP address
   - ❍ **D.** The master IP address

3. Which of the following first-hop redundancy protocols provides load-balancing capabilities from its group members?
   - ❍ **A.** HSRP
   - ❍ **B.** VRRP
   - ❍ **C.** GLBP
   - ❍ **D.** None of the above

## Answers

1. **B** is correct. VRRP is a standards-based FHRP that serves as an alternative to Cisco's proprietary HSRP.
2. **C** is correct. In HSRP, clients point to the virtual IP address, which serves as their default gateway.
3. **C** is correct. GLBP provides load-balancing capabilities, where each member of the group forwards traffic to the appropriate gateway.

# Multicast

Multicast communication is an efficient way of using network resources when moving network-intensive workloads such as video and audio. IP multicast is a bandwidth-conserving technology that can deliver a single stream of information to thousands of receivers simultaneously. Examples of applications that take advantage of multicast include video conferencing, distance learning, and distribution of software and news.

## Multicast Fundamentals

Traditional IP communication between hosts on a network is one of three types:

- **Unicast (one-to-one communication):** A host sends packets to another single host anywhere.
- **Broadcast (one-to-all communication):** A host sends packets to all hosts in the same broadcast domain.
- **Multicast (one-to-many communication):** A host sends packets to a subset of all hosts in the multicast domain.

IP multicast routing allows a host (source) to move packets to a group of hosts (receivers) within an IP network by using an IP multicast group address. The host that is sending traffic inserts the multicast group address in the IP destination field of the packet. The routers and multilayer switches between the source and receiver forward the incoming IP multicast packet out all interfaces that lead to members of the multicast group. Any host, regardless of whether it is a member of a group, can send to a group. However, only members of a specific group can receive the message.

Multicast uses two protocols to implement IP multicast routing:

- **Protocol Independent Multicast (PIM):** PIM is used between routers to build a multicast tree and track which multicast packets to forward to each other and to their locally connected LANs.
- **Internet Group Management Protocol (IGMP):** IGMP is used between hosts on a LAN and routers on that LAN to track which multicast groups hosts belong to.

# Multicast Group Addressing

A multicast group is identified by its multicast group address. Multicast packets are delivered to that multicast group address. A host must join the group that the address identifies to receive the data sent to a multicast address. The hosts that belong to a multicast group are referred to as *group members*. The data is sent to the multicast address and received by all the hosts that are part of that group. These hosts indicate that they wish to receive traffic sent to that group.

The Internet Assigned Numbers Authority (IANA) has assigned IP multicast addresses to the IPv4 Class D address space. Multicast host group addresses can be in the range 224.0.0.0 to 239.255.255.255. A multicast address is chosen and assigned at the source (sender) for the receivers in a multicast group.

Table 4.1 shows a summary of the multicast address ranges.

TABLE 4.1 **Multicast Address Ranges and Uses**

| Type of Addresses | Range | Description |
|---|---|---|
| Reserved link-local addresses | 224.0.0.0–224.0.0.255 | Reserved for network protocols on the local network segment |
| Globally scoped addresses | 224.0.1.0–238.255.255.255 | Reserved for sending multicast data between organizations and across the Internet |
| Source-specific multicast (SSM) addresses | 232.0.0.0–232.255.255.255 | Reserved for use with the SSM datagram delivery model, where data is forwarded only to receivers that have explicitly joined the multicast group |
| GLOP addresses | 233.0.0.0–233.255.255.255 | Reserved for statically defined addresses by organizations that already have an assigned autonomous system (AS) domain number |
| Limited-scope addresses | 239.0.0.0–239.255.255.255 | Reserved as administrative or limited-scope addresses for use in private multicast domains |

# Internet Group Management Protocol (IGMP)

IGMP is a protocol used to dynamically register individual hosts in a multicast group on a particular LAN. When you enable PIM on an interface, you are also enabling IGMP. IGMP provides a means of automatically controlling

and limiting the flow of multicast traffic throughout a network with the use of special multicast queriers and hosts:

- **Querier:** A querier is a network device that sends query messages to discover which network devices are members of a particular multicast group.
- **Host:** A host is a receiver, including a router, that sends report messages (in response to query messages) to inform the querier of host membership. Hosts use IGMP messages to join and leave multicast groups.

Hosts identify multicast group memberships by sending IGMP messages to their local multicast device. Under IGMP, devices listen to IGMP messages and periodically send out queries to discover which groups are active or inactive on a particular network.

**ExamAlert**

When preparing for the ENCOR exam, make sure you understand the features and differences between the three IGMP versions.

There are three versions of IGMP: IGMP version 1 (IGMPv1), IGMP version 2 (IGMPv2), and IGMP version 3 (IGMPv3). These versions are interoperable on Cisco Catalyst switches. Let us briefly look at the features of and differences between these versions:

- **IGMPv1:** IGMPv1, which is defined in RFC 1112, primarily uses a query/response model that enables the multicast router and multilayer switch to find which multicast groups are active (that is, have one or multiple hosts interested in a multicast group) on the local subnet.
- **IGMPv2:** IGMPv2, which is defined in RFC 2236, extends IGMP functionality by providing features such as the IGMP leave process to reduce leave latency, group-specific queries, and an explicit maximum query response time. IGMPv2 also adds the capability for routers to elect the IGMP querier without depending on the multicast protocol to perform this task.
- **IGMPv3:** IGMPv3, which is defined in RFC 3376, supports basic IGMPv3 snooping support (BISS), which includes support for the snooping features on IGMPv1 and IGMPv2 switches and IGMPv3 membership report messages. BISS restricts the flooding of multicast traffic when a network includes IGMPv3 hosts. It also supports SSM.

When a receiver wants to receive a multicast stream from a multicast source, it sends an unsolicited membership report, referred to as an *IGMP join*, to the local router for the group address that it intends to join (for example, 239.1.1.1). The local router then forwards this request up to the source, using a PIM join message. When the multicast stream starts coming in to the local router, it is forwarded downstream to the subnet where the receiver that made the request resides. Figure 4.9 illustrates this IGMP join process.

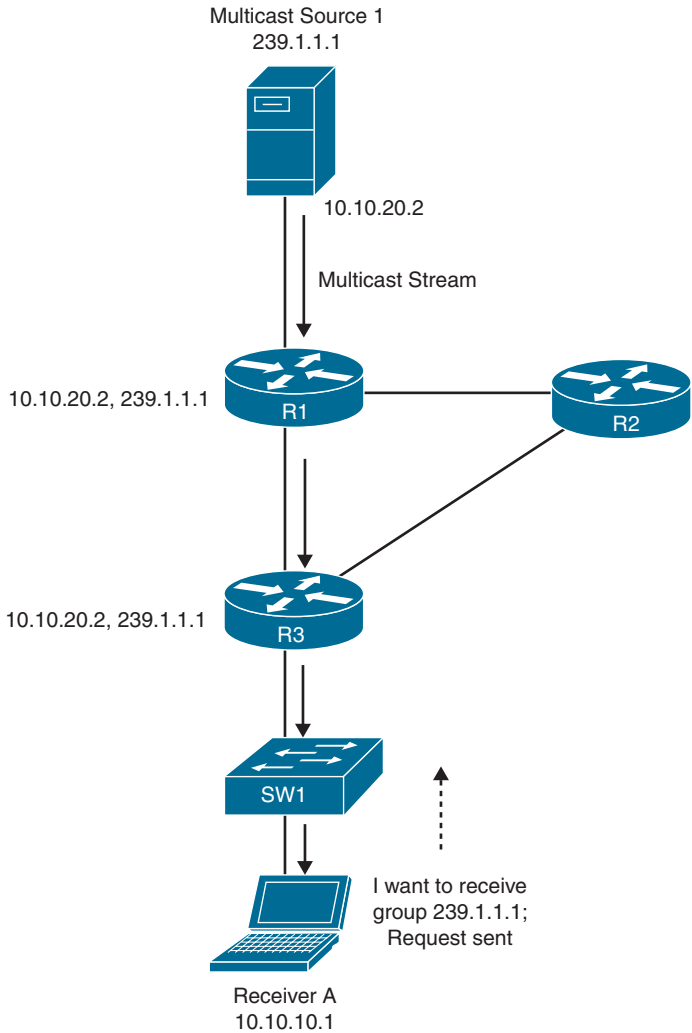



FIGURE 4.9 **The IGMP Join Process**

Let's look at the steps that IGMPv3 uses for both the join and the leave operations.

The following steps occur as part of the IGMPv3 join process:

1. When a host wants to join a group, it sends an IGMPv3 membership report to 224.0.0.22 with an empty EXCLUDE list.
2. When a host wants to join a specific channel, it sends an IGMPv3 membership report to 224.0.0.22 with the address of the particular source included in the INCLUDE list.
3. When a host wants to join a group, excluding particular sources, it sends an IGMPv3 membership report to 224.0.0.22, excluding those sources in the EXCLUDE list.

IGMPv3 enhances the leave process used by IGMPv2 by introducing the capability for a host to stop receiving traffic from a particular group, source, or channel in IGMP by including or excluding sources, groups, or channels in IGMPv3 membership reports.

For completeness, let us briefly look at the leave process of IGMPv2 that IGMPv3 builds on. IGMPv2 uses a leave-group message to indicate that it wishes to stop receiving multicast traffic for a specific group. When an IGMPv2 host leaves a multicast group, it sends a leave-group message to the all-devices multicast group if it was the last host to respond to a query with a membership report for that group.

To wrap up this part of our multicast discussion, let us look at Multicast Listener Discovery (MLD), which IPv6 hosts use. MLD is an IPv6 protocol that a host uses to request multicast data for a particular multicast group. Using the information obtained through MLD, the software maintains a list of multicast groups or channel memberships on a per-interface basis. The devices that receive MLD packets send the multicast data they receive for requested groups out the network segment of the known receivers.

MLDv1 is derived from IGMPv2, and MLDv2 is derived from IGMPv3. IGMP uses IP protocol 2 message types, while MLD uses IP protocol 58 message types, which are a subset of the ICMPv6 messages.

## IGMP Snooping

By default, Layer 2 switches treat multicast traffic the same as broadcast traffic: This traffic floods it out all ports except the port it came in on.

IGMP snooping, as defined in RFC 4541, examines the Layer 2 IP multicast traffic within a VLAN to discover the ports where interested receivers reside. Using the port information gathered, IGMP snooping can reduce bandwidth consumption in a multi-access LAN environment to avoid flooding the entire

VLAN. IGMP snooping tracks which ports are attached to multicast-capable routers to help the routers forward IGMP membership reports. IGMP snooping operates on IGMPv1, IGMPv2, and IGMPv3 control plane packets; Layer 3 control plane packets are intercepted and influence the Layer 2 forwarding behavior.

At a high level, IGMP snooping uses a switch to examine, or "snoop," Layer 3 information (IGMP join/leave messages) in the IGMP packets that are sent between hosts and a router. When a switch hears an IGMP host report from a host in a specific multicast group, the switch adds the host's port number to the multicast table entry. Similarly, when the switch hears an IGMP leave group message from a host, it removes the multicast table entry for that particular host.

# Protocol Independent Multicast (PIM)

To wrap up this chapter, let us look at PIM. Routers capable of moving multicast traffic create distribution trees that control the path that IP multicast traffic takes through the network in delivering traffic to all receivers. The two basic types of multicast distribution trees are source trees and shared trees:

- **Source trees:** A source tree is the simplest form of a multicast distribution tree, with its root at the source and branches forming a spanning tree through the network to the receivers. Since this tree uses the shortest path through the network, it is also referred to as the shortest path tree (SPT). The notation (S, G), pronounced "S comma G," enumerates an SPT, where S is the IP address of the source and G is the multicast group address. An example would be (192.168.1.100, 224.1.1.1). Using the SPT notation, a separate SPT exists for each sender to each group.
- **Shared trees:** Unlike source trees that have their root at the source, shared trees use a single common root placed at some chosen point in the network. This shared root is known as a rendezvous point (RP). Source traffic is sent toward the RP on a source tree. The traffic is then forwarded down the shared tree from the RP to reach all of the receivers.

PIM uses unicast routing information to create a distribution tree along the reverse path from the receivers toward the multicast source. The multicast routers in the path then forward packets along the distribution tree from the source to the receivers. Reverse-path forwarding (RPF) is a concept in multicast forwarding that enables routers to forward multicast traffic down the distribution tree correctly. RPF makes use of the existing unicast routing table to determine the upstream and downstream neighbors. A router forwards a

multicast packet only if it is received on the upstream interface. This RPF check helps guarantee that the distribution tree will be loop free.

When a multicast packet arrives at a router, the router performs an RPF check on the packet. If the RPF check succeeds, the packet is forwarded. Otherwise, it is dropped.

PIM is routing protocol independent and can leverage whichever unicast routing protocols are used to populate the unicast routing table (for example, EIGRP, OSPF, BGP) and static routes. Although PIM is called a multicast routing protocol, it uses the unicast routing table to perform the RPF check function instead of building up a completely independent multicast routing table. PIM does not send and receive routing updates between routers, as routing protocols do. PIM forwarding modes include the following:

- **PIM Dense Mode (PIM-DM):** PIM-DM uses a push model to initially flood multicast traffic throughout the network. The push model is an aggressive method for delivering data to the receivers. This method is efficient in specific deployments in which there are active receivers on every subnet in the network. Routers that have no downstream neighbors prune back the unwanted traffic. PIM-DM supports only source trees—that is, (S, G) entries—and cannot be used to build a shared distribution tree.
- **PIM Sparse Mode (PIM-SM):** PIM-SM uses a pull model to deliver multicast traffic. PIM-SM distributes information about active sources by forwarding data packets on the shared tree. Only network segments with active receivers that have explicitly requested the data receive the traffic. Because PIM-SM uses shared trees initially, it requires the use of an RP.
- **PIM Sparse-Dense Mode:** In this mode, the router handles both dense groups and sparse groups at the same time.
- **Bidirectional PIM (Bidir-PIM):** Bidir-PIM is an enhancement of the PIM protocol that is designed for efficient many-to-many communications within a PIM domain. Multicast groups in bidirectional mode can scale to an arbitrary number of sources with only a minimal amount of additional overhead.
- **Source-Specific Multicast (SSM):** SSM is an extension of the PIM protocol that provides an efficient data delivery mechanism in one-to-many communications. SSM enables a receiving client, once it has learned about a particular multicast source through a directory service, to receive content directly from the source rather than receiving it using a shared RP.

ExamAlert

For the ENCOR exam, make sure you understand the purpose of an RP and when it is used.

Let us look a little more closely at RPs. A rendezvous point is a role that a network device performs when operating in PIM-SM. It is required only in networks running PIM-SM. An RP acts as the meeting place for sources and receivers of multicast data. In a PIM-SM network, sources must send their traffic to the RP. The traffic is then forwarded to receivers down a shared distribution tree. By default, the RP is needed only to start new sessions with sources and receivers. Thus, the RP experiences little overhead from traffic flow or processing.

In PIM-SM, it is mandatory to choose one or more routers to act as the RP. Related to group-to-RP mapping, an RP can be configured statically in the multicast domain or dynamically by configuring Auto-RP or PIM Bootstrap Router:

- **Static RP:** With static RP, you statically configure an RP for a multicast group range on every router in the multicast domain. It is simple to configure and can be ideal in small network environments. However, it can introduce administrative overhead in larger environments. If an RP fails, there is no failover method to take over the function of the failed RP. Another downside of statically configuring RP is that if several RPs are active for different groups, information about which RP is handling which group must be known to all routers—and this involves additional configuration.
- **Auto-RP:** This is a Cisco-proprietary method that automates the distribution of group-to-RP mappings in a PIM network. Auto-RP provides advantages like load splitting when using different RPs and makes it possible to have different RPs serve different group ranges or serve each other as backups. Auto-RP operates by using two main components:
  - **Candidate RPs:** These RPs advertise their willingness to become RPs by sending RP announcement messages at 60-second intervals to the well-known multicast group address 224.0.1.39 (CISCO-RP-ANNOUNCE).
  - **RP mapping agents:** These agents receive the RP announcement messages from the RPs and arbitrate conflicts. RP mapping agents join group 224.0.1.39 to receive RP announcements and advertise the mappings to 224.0.1.40 (CISCO-RP-DISCOVERY) every 60 seconds

by default (or when changes occur). The RP mapping agents then send the consistent group-to-IP address mappings to all other devices.

- **PIM Bootstrap Router (BSR):** BSR is a nonproprietary method defined in RFC 5059 that is used to discover and announce RP-set information for each group prefix to all the routers in a PIM domain. After the election of the BSR, candidate RPs use unicast to announce to the BSR their willingness to be the RP. The BSR advertises the entire group-to-RP mapping to the router's link-local address 224.0.0.13. Unlike with the RP mapping agent in Auto-RP, every router in the BSR network is responsible for selecting the RP.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Multicast uses which of the following protocols to implement IP multicast routing? (Choose two.)
   - ❍ **A.** PIM
   - ❍ **B.** OSPF
   - ❍ **C.** MP-BGP
   - ❍ **D.** IGMP

2. Which PIM forwarding mode was designed for efficient many-to-many communications?
   - ❍ **A.** PIM-DM
   - ❍ **B.** PIM-SM
   - ❍ **C.** Bidir-PIM
   - ❍ **D.** SSM

## Answers

1. **A** and **D** are correct. PIM and IGMP are critical to implementing IP multicast routing.
2. **C** is correct. Bidir-PIM is an enhancement of the PIM protocol that is designed for efficient many-to-many communications within a PIM domain.

# Review Questions

1. Which of the following protocols and ports does NTP use?
   - ❍ **A.** UDP 123
   - ❍ **B.** TCP 123
   - ❍ **C.** UDP 119
   - ❍ **D.** TCP 119

2. Which of the following commands is used to view the NAT translation table of a router?
   - ❍ **A. show nat translations**
   - ❍ **B. show ip nat translations**
   - ❍ **C. show nat**
   - ❍ **D. show ip translations**

3. True or false: In VRRP, the device with the lowest priority becomes the master.
   - ❍ **A.** True
   - ❍ **B.** False

4. Which of the following group-to-RP mapping technologies is Cisco proprietary?
   - ❍ **A.** Static RP
   - ❍ **B.** SSM
   - ❍ **C.** BSR
   - ❍ **D.** Auto-RP

# Answers to Review Questions

1. **A** is correct. NTP uses UDP as its transport protocol, operating on port 123.
2. **B** is correct. Verification of the NAT translation table is done using the **show ip nat translations** command.
3. **B** is correct. With VRRP, the device with the highest priority becomes the master, and the priority can range from 0 to 255.
4. **D** is correct. Auto-RP is a Cisco-proprietary method that automates the distribution of group-to-RP mappings in a PIM network.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*
- *CCNP and CCIE Enterprise Core & CCNP Advanced Routing Portable Command Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers enterprise wireless.

CHAPTER 5

# Enterprise Wireless

**This chapter covers the following official ENCOR 350-401 exam objectives:**

- 3.3 Wireless
- 3.3.a Describe Layer 1 concepts, such as RF power, RSSI, SNR, interference noise, band and channels, and wireless client devices capabilities
- 3.3.b Describe AP modes and antenna types
- 3.3.c Describe access point discovery and join process (discovery algorithms, WLC selection process)
- 3.3.d Describe the main principles and use cases for Layer 2 and Layer 3 roaming
- 3.3.e Troubleshoot WLAN configuration and wireless client connectivity issues

This chapter is divided into three sections. It starts by covering the basic theory behind wireless Layer 1 concepts and how data is sent over the air. Then, it looks at radio frequency (RF) power, received signal strength indicator (RSSI), signal-to-noise ratio (SNR), interference noise, band and channels, and wireless client devices capabilities. The second section looks at enterprise wireless infrastructure and how to build topologies with multiple access points (APs). This section looks at the various AP modes and antenna types used in enterprise wireless deployments. It also talks about how lightweight APs discover and join a wireless LAN controller (WLC). The final section of this chapter covers the principles that facilitate the roaming of wireless clients between APs. It examines both Layer 2 and Layer 3 roaming. This section also covers troubleshooting of WLAN configuration and wireless client connectivity issues.

**This chapter covers the following technology topics:**

- Wireless Basics
- WLC and AP Operation and Pairing
- Wireless Roaming
  - Troubleshooting WLAN Configuration and Client Connectivity Issues

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. What do wireless client receivers use to measure the received signal power level?
2. In which mode does an AP passively monitor the WLAN environment over a specifically configured channel?
3. Which AP mode facilitates simplified management and is ideal for environments where a large number of APs are required in a WLAN deployment?
4. Which type of antenna is designed to provide a 360-degree radiation pattern?
5. When multiple wireless clients are encountering connectivity issues in the same area, where would be a good starting point for troubleshooting?

### Answers

1. Received signal strength indicator (RSSI)
2. Sniffer
3. Lightweight mode
4. Omnidirectional antenna
5. On the access point

# Wireless Basics

This section starts by looking at the basics of how data is moved wirelessly between devices on a network at Layer 1, using radio frequency (RF) waves. Throughout this chapter, the terms *wireless LAN controller (WLC)* and *controller* are used interchangeably.

# Radio Frequency (RF)

A wireless LAN transmits a wireless signal over radio frequency (RF) waves to move data from one wireless device to another. A radio wave is an electromagnetic field (EMF) radiated away from a transmitter. The electric and magnetic field transmitted away from the transmitter propagates out as traveling waves, and a receiver receives this energy. The signal sent from the transmitter

continues alternating—or changing by being cycled up and down—and this keeps the electric and magnetic fields cycling and moving outward.

You can measure the waves that are transmitted from a transmitter by looking at the property of the waves or how often the signal makes one up and down cycle in 1 second. This measurement is done in hertz (Hz), which is a frequency unit that is 1 cycle per second. Radio frequency can range between 3 kHz and 300 GHz. Technologies such as low-frequency radios, microwaves, televisions, and radar all operate within this range. The frequency range 2.4 to 5 GHz is used for WLAN communications.

A range of frequencies can be used for the same purpose, and such a range is often referred to as a *band* of frequency. For example, AM radio stations use the range 530 kHz to 1710 kHz, which is appropriately referred to as the *AM band*. Wireless LAN communication uses two main frequency ranges, 2.400–2.4835 GHz and 5.150–5.825 GHz, which are referred to as the 2.4 GHz band and the 5 GHz band, respectively. The 5 GHz band, which is more commonly used of these two bands, takes advantage of four distinct bands for communication:

- 5.150–5.250 GHz
- 5.250–5.350 GHz
- 5.470–5.725 GHz
- 5.725–5.825 GHz

Further, to keep wireless communication orderly, bands are divided into various distinct channels that are numbered and assigned to specific frequencies. Channels are spaced at regular intervals, and this channel spacing is known as the *channel separation* or *channel width*. RF can inadvertently occupy neighboring frequencies because the RF signal is not narrow, and it spills above and below a center frequency. The frequency range that is needed to transmit a signal is known as the *signal bandwidth*. In WLAN communication, the signal bandwidth is defined as part of a wireless standard. (The various standards are covered later in this chapter.)

The signal bandwidth should preferably be less than the channel width. This allows for different signals to be transmitted on every possible channel, with no possibility of two signals overlapping and interfering with each other. At times, signals must be placed on more distant channels to prevent overlap, but this placement limits the number of channels that can be used in the band.

The strength of the RF waveform is measured as amplitude, which is basically the height from the top peak to the bottom peak of the waveform. This signal strength is a measure of power, in watts (W). A WLAN transmitter typically

has a signal strength between 0.1 W (100 mW) and 0.001 W (1 mW). To measure the amount of change in power at any discrete point, we use terminology like decibel (dB), decibel isotropic (dBi), and decibel milliwatts (dBm).

The term dB is usually used for attenuation and amplification of the power level. It is a logarithmic ratio of the signal to another standardized value. For example, in dBm, the value is being compared to 1 milliwatt, and in dBw, the value is being compared to 1 watt.

On the other hand, the term dBi is used to describe the power gain rating of a wireless antenna. For example, if an antenna has a gain of 2.2 dBi, it means the maximum energy density of the antenna is 2.2 dB greater than that of an isotropic antenna. All Federal Communications Commission (FCC) calculations use dBi as the measurement.

The term dBm uses the same calculations as dB but references a value of 1 milliwatt. For example, if the power level at a radio jumps from 1 mW to 100 mW, the power level jumps from 0 dBm to 20 dBm. dBm can also be used to describe receiver sensitivity. Receiver sensitivity is measured in negative dBm (–dBm) because the signal decreases in value from its point of transmission to the receiver. The sensitivity value indicates the lowest power the receiver can receive before it considers the signal unintelligible.

Even though the transmitted power based on the radio setting is rated in either dBm or watts, the maximum energy density from an antenna is measured as effective isotropic radiated power (EIRP). EIRP is the summation of the dB values of the various components. It is calculated by adding the transmitted power (in dBm) to antenna gain (in dBi) and subtracting cable loss (in dB).

As it relates to RF waves moving data, RF is called a *carrier signal* because it is used to carry meaningful information. In the case of AM and FM radios, it is used to transport audio signals. In terms of TV, the carrier signal transports both audio and video. And for the purposes of this chapter, a wireless LAN carrier signal carries data. The carrier signal needs to be altered in a way that indicates the information to be transported. This is known as *modulation*. The receiver demodulates, or interprets, the information based on the changes in the carrier signal.

**ExamAlert**

It is important to understand and be able to differentiate between free space path loss, received signal strength indicator, and signal-to-noise-ratio for the ENCOR exam.

# Free Space Path Loss

As an RF signal is transmitted from a transmitter to a receiver and travels through free space, its amplitude decreases. Antenna gain, cable loss, data rate, link distance, transmitter power, receiver sensitivity, and other variables play roles in determining how far the RF signal moves from the transmitter. Even if there is no obstacle in the path between the transmitter and the receiver, the signal degrades as it travels through free space. This is known as the *free space path loss*. It is an exponential function, meaning that the signal falls off quickly near the transmitter but more slowly as it is transmitted further away. In other words, as the client/receiver moves away from the transmitter, the signal gets weaker.

# Received Signal Strength Indicator (RSSI)

Two factors determine how much signal a receiver gets. The first is received signal strength indicator (RSSI). The receiver typically measures the received signal power level by using the RSSI scale. Doing this calculation is not an easy task because the receiver has no way of determining how much power was initially transmitted.

RSSI can be measured on a scale of 0 to 255, as defined by the 802.11 standards. However, different vendors can define their own scale within 0–255. For the RSSI grade value, there is an equivalent dBm value, and it varies with vendors. The main difference really is that RSSI is a relative index, and dBm is an absolute number representing the power level, in milliwatts.

# Signal-to-Noise Ratio (SNR)

Wireless devices need to distinguish between legitimate signals that they should be listening to and background signals on the spectrum. This concept is known as the *signal-to-noise ratio (SNR)*. SNR is the difference between a wireless signal that is received and the *noise floor*—that is, background transmissions emitted from devices that are too far away to be intelligible or from devices that are inadvertently creating interference on the same frequency. SNR is measured as positive value between 0 dB and 120 dB. The closer the value is to 120 dB, the better. For example, if a client device receives a signal of –70 dBm and the floor noise is –90 dBm, then effective SNR is 20 dB. SNR is calculated as follows:

Signal Strength – Noise Floor = SNR

For example:

–70 dBm – –90 dBm = 20 dB

A higher SNR value is better because the further a received signal is from the noise floor, the better the signal quality. Signals close to the noise floor can be susceptible to data corruption, requiring retransmission between the transmitter and the receiver. Such retransmission degrades the wireless throughput and latency.

# IEEE Wireless Standards

IEEE 802.11 is the working group within the Institute of Electrical and Electronics Engineers (IEEE) that is responsible for wireless LAN standards at the physical and link layers. Over the years, multiple implementations of the 802.11 standards have been developed. Let us take a brief look at them:

- **802.11:** This is the original standard, which is now obsolete. There were two variations of 802.11, both of which operated at a speed of up to 2 Mbps on an RF of 2.4 GHz.
- **802.11a:** This standard operates with a speed of up to 54 Mbps on an RF of 5 GHz. However, realistically, due to error correction code, the actual throughput was much lower. 802.11a is incompatible with the 802.11b and 802.11g wireless standards. 802.11a signals cannot penetrate as far as 802.11b signals because they are easily absorbed by walls and other solid objects in the signal path due to the smaller wavelength. 802.11a also suffers significantly from interference.
- **802.11b:** This standard operates with a speed of up to 11 Mbps on an RF of 2.4 GHz. It provides a more extended range than 802.11a and can better penetrate solid objects. Devices using 802.11b at 2.4 GHz, including baby monitors, microwaves, and cordless phones, can interfere with 802.11b wireless signals.
- **802.11g:** This standard operates with a speed of up to 54 Mbps on an RF of 2.4 GHz. 802.11g is backward compatible with 802.11b, though with reduced throughput. As with 802.11b, devices operating in the 802.11b 2.4 GHz band can also interfere with wireless LAN communication.
- **802.11n (Wi-Fi 4):** This standard operates with a speed of 54 to 600 Mbps in both 2.4 and 5 GHz bands (typically referred to by vendors as *dual band*).

802.11n is backward compatible with 802.11a/b/g devices, though with limited data rates. The 802.11n standard added support for multi-input, multi-out (MIMO) technology.

- **802.11ac (Wi-Fi 5):** This standard operates with a speed of 450 Mbps to 1300 Mbps (1.3 Gbps) using MIMO technology in the 5 GHz band. 802.11ac is backward compatible with 802.11a/n. Compared to 802.11n, 802.11ac includes wider channels (80 or 160 MHz as opposed to 40 MHz) and more spatial streams (eight versus four).
- **802.11ax (Wi-Fi 6):** This standard operates with a theoretical speed of up to 9.6 Gbps using MIMO technology in the 2.4 and 5 GHz bands. Wi-Fi 6E allows for the expansion of 802.11ax into the 6 GHz band. 802.11ax is typically deployed in dense environments (network with various or many WLAN clients).

## Multiple Radios

To support higher speeds, some 802.11 standards require that APs and clients have multiple antennas supporting MIMO technology. MIMO works by using multiple antennas at both the transmitter and the receiver ends to improve the performance of the wireless communication. The use of multiple radio components creating multiple radio chains gives the 802.11n, 802.11ac, and 802.11ax better performances. For example, a 2×3 MIMO system would consist of three radio chains with two transmitters and three receivers. Prior to 802.11n, wireless devices used a single transmitter and a single receiver. Such a system is known as a *single-in, single-out (SISO)* system. MIMO encompasses three technologies that help to scale performance:

- **Spatial multiplexing:** For increased throughput, data communication over a wireless LAN can be multiplexed across two or more radio chains. These chains all operate on the same channel but are separated by spatial diversity. This is referred to as *spatial multiplexing*. The number of spatial streams supported is usually designated with a colon, followed by the number of unique spatial streams. For example, 3×3:2 would indicate three transmitters, three receivers, and two unique spatial streams.
- **Transmit beamforming:** With a single radio chain, when a transmitter sends an RF signal, any of the receivers have the same opportunity to receive and interpret the signal; that is, the transmitter does nothing to prefer one receiver over another. With MIMO, the same signals from

the transmitter, using multiple antennas, travel to the client location more efficiently. With transmit beamforming, you can improve signal quality and SNR by altering the phase of the signal as it is being fed into the transmitting antenna so that the RF signal can arrive in phase at a receiver.

- **Maximum-ratio combining:** As you saw earlier in this chapter, with MIMO, the same signal can be transmitted over multiple antennas. A receiver can use multiple antennas and radio chains to receive multiple copies of the signal. One copy of the signal might be better than the others, or it might be better for a period and then become worse than the others. Maximum-ratio combining can combine the multiple copies of the signal to present the best one at a particular time. The end result of maximum-ratio is a reconstructed signal with an improved SNR.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. What unit does the FCC use to refer to the power gain rating of a wireless antenna?
   - ❍ **A.** mW
   - ❍ **B.** dBw
   - ❍ **C.** dBi
   - ❍ **D.** dBm

2. Although vendors can use a different range, what is the range used to measure RSSI as defined in the 802.11 standards?
   - ❍ **A.** 0 to 255
   - ❍ **B.** 0 to 100
   - ❍ **C.** 0 to 90
   - ❍ **D.** 0 to 70

3. If a client device receives a –70 dBm signal, and the floor noise is –90 dBm, what is the effective SNR?
   - ❍ **A.** 180 dB
   - ❍ **B.** –20 dB
   - ❍ **C.** 20 dB
   - ❍ **D.** 40 dB

# Answers

1. **C** is correct. The unit dBi is used to describe the power gain rating of a wireless antenna. All FCC calculations use dBi as the measurement.
2. **A** is correct. RSSI can be measured on a scale of 0 to 255, as defined by the 802.11 standards. However, different vendors can define their own scales within 0–255.
3. **C** is correct. The effective SNR in this case is 20 dB.

# WLC and AP Operation and Pairing

This section looks at the various AP modes as well as the different AP types. It also covers how lightweight APs discover and join a wireless LAN controller. Cisco APs operate in one of two modes: autonomous mode or lightweight mode. Chapter 20, "Wireless LAN Deployments Deployments," covers these AP modes to a small extent, along with the various WLAN deployment models for them.

**ExamAlert**

For the ENCOR exam, you should understand the two operational modes of an AP.

Let us take a closer look at these AP modes:

- **Autonomous mode:** An AP that is operating in autonomous mode is self-contained, containing both wired and wireless hardware to bridge the wireless client's SSID to the wired VLAN infrastructure at the access layer. On a standalone AP, one or more SSIDs can be provisioned, each mapping to the respective VLANs on the wired infrastructure. For multiple SSIDs and, therefore, multiple VLANs, the ports need to be trunked all the way from the distribution layer, where routing between VLANs occurs, down to the access layer switches that connect the APs, and then finally all the way down to the APs. All of the autonomous APs must be managed individually using Telnet, SSH, or a web interface unless you are using a platform like Cisco Prime Infrastructure or Cisco DNA Center. As you can imagine, scaling the wireless infrastructure using autonomous APs can be difficult as the infrastructure grows.
- **Lightweight mode:** An AP that is operating in lightweight mode uses Lightweight Access Point Protocol (LWAPP) or Control and Provision of Wireless Access Points (CAPWAP) to allow the WLC to communicate with an AP. Lightweight APs are ideal for environments where a large number of APs are required in a WLAN deployment. As you scale the number of APs, it is simpler to manage devices in lightweight mode because each AP is configured automatically and managed by the WLC. Further details related to the configuration of lightweight APs in a centralized deployment model are provided in Chapter 20.

ExamAlert

Knowing the different modes that you can use in deploying APs would be helpful for the ENCOR exam.

You can deploy Cisco APs in various modes:

- **Local:** Local mode is the default mode for lightweight APs. In this mode, the AP maintains a tunnel back to the WLC, and client traffic is centrally switched on the WLC. If the AP loses its connection to the WLC, the AP stops forwarding traffic until it joins another WLC. When the AP is not transmitting data, it can be scanning other channels for noise level, measuring interference, or discovering rogue devices.
- **Monitor:** An AP in monitor mode does not transmit any data. Instead, the AP receiver acts as a dedicated sensor. It scans for rogue APs, checks for IDS events, and determines the client's position through location-based services.
- **FlexConnect:** An AP in FlexConnect mode is basically a local AP deployed at a remote site establishing a CAPWAP tunnel back to a headquarter's WLC. If the WAN between the remote site and the headquarters (AP and WLC) is down, the AP can switch traffic between a VLAN and SSID.
- **Bridge:** An AP in bridge mode becomes a dedicated bridge between two networks. Two APs in bridge mode can be used to link multiple locations. Multiple APs can be connected to form an indoor or outdoor mesh network.
- **Flex+Bridge:** For an AP in this mode, FlexConnect is enabled on a mesh AP.
- **SE-Connect:** An AP in SE-Connect mode allows connection to an AP using Cisco Spectrum Expert to gather information about the RF spectrum surrounding an AP. This mode is used strictly for troubleshooting.
- **Sniffer:** An AP in sniffer mode is used for troubleshooting purposes. In passive mode, an AP passively monitors the WLAN environment over a specifically configured channel. The data is then tunneled to an endpoint running a protocol analysis tool, such as Wireshark, to analyze packets.
- **Rogue detector:** An AP in rogue detector mode does not support wireless clients. Instead, the AP scans for MAC addresses, which it hears on the wired and wireless networks. Rogue devices would be devices found on both networks.

# AP and WLC Interaction

Next, let's look at the process that an AP goes through to join a WLC. There are three broad ways an AP can discover a WLC:

- By using prior knowledge of the WLC
- By having DHCP and DNS preprovisioned information about the WLC
- By broadcasting on the local subnet to find the WLC

Generally, an AP sends a unicast CAPWAP discovery request to the WLC interface over UDP port 5246 or broadcasts on the local segment. If a WLC exists, it responds with a CAPWAP discovery response to the AP. Let's look more closely at these discovery steps:

1. The AP boots up and receives an IP address from a DHCP server if such an address was not statically assigned previously.
2. The AP sends discovery requests to WLCs to learn their management addresses via:
   - DHCP option 43 (suitable for when the AP and the WLC are on separate subnets)
   - A DNS entry for cisco-capwap-controller
   - Management IP addresses of WLCs that the AP remembers
   - A Layer 3 broadcast on the subnet (where the AP automatically looks on the local subnet for controllers with a 255.255.255.255 local broadcast)
   - Statically configured information
   - Controllers present in the mobility group of the WLC that the AP last joined
3. The AP sends a discovery request to every WLC on the list and waits for the WLC's discovery reply, which contains the system name, the AP-manager IP addresses, the number of APs already attached to each AP-manager interface, and overall excess capacity for the controller.
4. If the AP receives a discovery reply from the WLC, the AP looks at the WLC list and sends a join request to a WLC using the following order:
   - Primary controller system name
   - Secondary controller system name

- Tertiary controller system name
- Master controller (if the AP has not been previously configured with any primary, secondary, or tertiary controller names)
- Load balancing across WLCs using the excess capacity value in the discovery response (if none of the above are seen)

5. If the AP does not receive a join response from its choice, the AP tries the next WLC in the list unless the WLC is a configured controller (primary/secondary/tertiary). (Join requests must have a valid certificate to get a join response from a WLC.)
6. When it receives the join reply, the AP checks to make sure it has the same image as the WLC. If it doesn't, the AP downloads the image from the controller and reboots to load the new image that it downloaded, and it starts the process over again from step 1.
7. If the AP has the same software image, it asks for the configuration from the WLC and moves into the registered state on the WLC.
8. After the AP downloads the configuration, it might reload again to apply the new configuration.

From the list of WLC discovery methods, the easiest method for deployment is to have the APs on the same subnet as the management interface of the WLC and let the AP broadcast on the segment find the WLC. The next easiest method of deployment is to use a DNS entry with DHCP. You can have multiple entries with the same DNS name. This allows the AP to discover multiple WLCs. Finally, if you cannot use either of these methods, you can statically configure the information necessary to join a WLC via the console port and the APs CLI.

An AP can discover multiple WLCs and not just the one that it chooses to join. When an AP joins a WLC, it uses keepalive messages at regular intervals to track the status of the WLC. If the WLC that it joins fails, the AP selects the next least-loaded WLC and joins it. However, as simple as it sounds, this process may not be ideal when a WLC has a larger number of APs and their associated clients connected to it. While APs are being joined to another WLC, the wireless clients are left stranded with no connectivity. Deploying multiple WLCs in a high-availability configuration using stateful switchover (SSO) helps to alleviate the issue of an AP having to join a new WLC in the event of a WLC failure.

SSO groups the WLCs in a high-availability pair where one of the WLCs takes on the active role, and the other takes on the standby role. The active controller keeps track of the CAPWAP tunnels, AP states, client states, confirmation, and image files and syncs them with the standby controller. Should the active controller fail, the standby controller already has the current state information for the APs and clients, making the failover process transparent.

To wrap up this section, let's look at the various antenna types that you can leverage for enterprise wireless deployments. The following terms will give you a better understanding of antenna types:

**ExamAlert**

For the ENCOR exam, you need to have a good understanding of antenna types and placement to maximize radio coverage.

- **Gain:** The gain of an antenna is a measure of its power, or how effectively the antenna can focus RF power in a certain direction.
- **Direction:** The direction is the shape of the transmission pattern.
- **Polarization:** Polarization is the electrical field wave's orientation with respect to the horizon.

Before getting further into the discussion of antennas, let us briefly look at plane patterns. This will help you solidify your understanding of both omnidirectional and directional antennas.

All antennas transmit power in a distinct three-dimensional shape radiating out of the antenna. These patterns can be conceptualized using graph patterns that are known as *radiation patterns*. A graph shows the shape of a radiation pattern by measuring antenna gain at one or multiple frequencies. This is done by taking cross-sections from different angles (that is, based on *patterns*).

There are basically two types of plane patterns:

- **Azimuth:** This is a bird's-eye view of the pattern of the antenna that shows gain reaching out on the horizon.
- **Elevation:** Elevation is a cross-section of an antenna radiation pattern at eye level with the AP, from an angle on the horizon.

Now that we have examined radiation patterns, let us take a look at antenna types.

**ExamAlert**

For the ENCOR exam, you need to understand the two antenna types and their differences.

Cisco offers two antenna types, each offering different coverage capabilities. As the gain of an antenna increases, there is a trade-off in its coverage area. Usually, high-gain antennas provide a longer coverage area but in a certain direction. Let us now look at the two types of antennas:

- **Omnidirectional antennas:** Omnidirectional antennas are designed to provide a 360-degree radiation pattern. They are typically used when coverage from an antenna in all directions is required. An omnidirectional antenna is usually shaped like a thin cylinder. It propagates the signal in all directions away from the cylinder—but not along the cylinder length.

  A common type of omnidirectional antenna is the dipole. Dipoles have two separate wires that radiate signals away when alternating current is applied across them. They usually have a gain of +2 to +5 dBi. Figure 5.1 shows Cisco Aironet dipole antennas.

FIGURE 5.1 **Cisco Aironet Dipole Antennas**

Many of Cisco's APs integrate the antennas within the device case so that they are not visible. Figure 5.2 shows a Cisco Catalyst 9120AX AP, which has two integrated omnidirectional antennas hidden inside.

FIGURE 5.2 **Cisco Catalyst AP with Integrated Omnidirectional Antennas**

Figure 5.3 shows the azimuth and elevation patterns of a Cisco C9105AXI AP omnidirectional antenna.

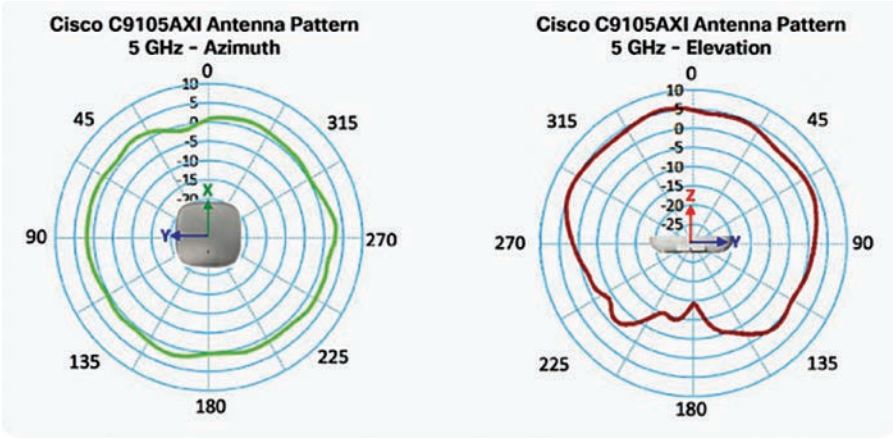



FIGURE 5.3 **Azimuth and Elevation Patterns of a Cisco C9105AXI AP Omnidirectional Antenna**

- **Directional antennas:** A directional antenna directs the energy it receives from the transceiver. By redirecting this energy, the antenna can focus more energy in one direction and less energy in another. With directional antennas, as the gain of the energy increases, the angle of radiation usually decreases, providing a greater coverage distance but a reduced coverage angle. These antennas are typically used indoors in elongated areas such as aisles or hallways. They are also ideal for outdoor applications away from buildings or for covering long distances between buildings.

A common type of directional antenna is a patch antenna, which produces a broad egg-shaped RF signal pattern that extends out from the flat patch surface. (Figure 5.4 shows a Cisco Aironet directional patch antenna.) Another common type of directional antenna is a Yagi antenna, which produces a more focused egged-shaped pattern that extends along the length of the antenna. Another common form of directional antenna is the parabolic dish antenna. As RF waves arrive from the line of sight, these waves reflect onto the center element on the antenna that faces the dish.

FIGURE 5.4 **Cisco Aironet Directional Patch Antenna**

Figure 5.5 shows the azimuth and elevation patterns of a Cisco C9124AXD AP directional antenna.

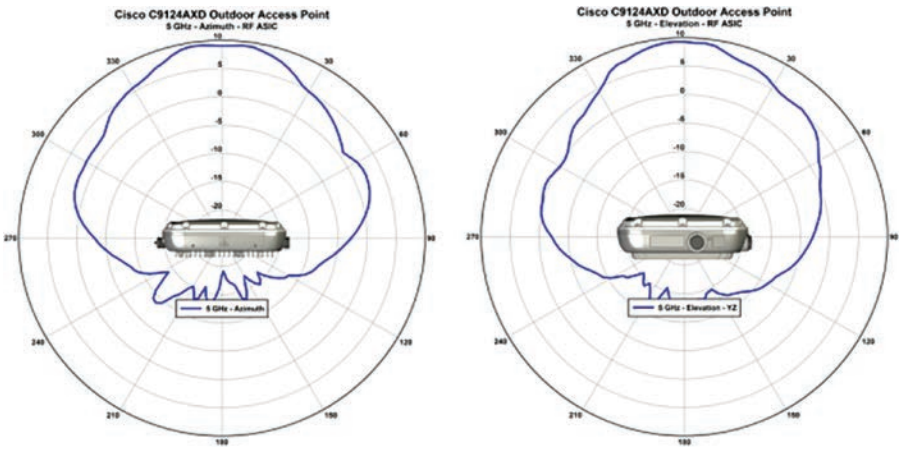



FIGURE 5.5 **Azimuth and Elevation Patterns of a Cisco C9124AXD AP Directional Antenna**

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. What is the default mode of operation for a lightweight AP?
   - ❍ **A.** Local
   - ❍ **B.** Monitor
   - ❍ **C.** Bridge
   - ❍ **D.** SE-Connect

2. What is the easiest way for a lightweight AP to discover a WLC?
   - ❍ **A.** Deploying a DNS entry with DHCP
   - ❍ **B.** Statically configuring an AP
   - ❍ **C.** Deploying through a mobility group
   - ❍ **D.** Placing an AP on the same subnet as the WLC management interface

3. True or false: An omnidirectional antenna is suitable when coverage from an antenna in all directions is required.
   - ❍ **A.** True
   - ❍ **B.** False

## Answers

1. **A** is correct. Local mode is the default mode for lightweight APs. In local mode, the AP maintains a tunnel back to the WLC, and client traffic is centrally switched on the WLC.
2. **D** is correct. The easiest method for deployment is to have the APs on the same subnet as the management interface of the WLC and let the AP broadcast on the segment to find the WLC.
3. **A** is correct. Omnidirectional antennas are designed to provide a 360-degrees radiation pattern and are typically used when coverage from an antenna in all directions is required.

# Wireless Roaming

This section looks at how wireless clients can roam from one AP/WLC pair to another. A wireless client may decide to roam from one AP to another for a number of reasons. One common reason for roaming is degradation of the wireless client's connection to the current AP. With roaming, there is some impact on the wireless client's connection as it roams between APs. This impact is due to the wireless client's scanning other channels for other APs, reassociating, and authenticating to the new AP.

Before getting into the various roaming scenarios, let us briefly look at some other common reasons that would cause wireless clients to roam:

- **Low RSSI:** When the wireless signal drops below a certain threshold, a wireless client may decide to roam. The roam trigger does not require active client traffic to initiate roaming.
- **Low SNR:** When the difference between the received signal strength and the noise floor drops below a threshold, a client device may decide to roam. The roam trigger does not require active client traffic in order to initiate a roaming.
- **Maximum data retry count exceeded:** An excessive number of retries can cause a wireless client to roam.
- **Proprietary load-balancing schemes:** Some wireless deployments have schemes that allow clients to roam to more evenly balance wireless client traffic across multiple APs. The roaming may be triggered in the WLAN infrastructure and then communicated to wireless clients using vendor-specific protocols.

Next, we will look at various roaming scenarios:

- **Roaming between autonomous APs:** A wireless client typically continuously scans the quality of a wireless connection, whether it is moving around or not. As the signal degrades for a wireless client, the client selects another AP, where it has a better signal, and reassociates with it. With APs operating in autonomous mode, each AP maintains its own table of associated clients.

  Say that a wireless client is connected to AP 1, and the signal is degraded. The client has a stronger signal from AP 2, so it decides to roam and reassociate with AP 2. Both AP 1 and AP 2 maintain their lists of associated wireless clients. After the roaming operation, both APs update their wireless client lists. If there are any leftover frames destined for the wireless

client after the roaming operation, they are forwarded over the wired network to AP 2 because this is where the MAC address for that wireless client that roamed now resides.

- **Intra-controller roaming:** The roaming process when using a WLC is similar to wireless clients roaming between autonomous APs because wireless clients have to reassociate to a new AP as they move around. However, with intra-controller roaming, because of the split-MAC architecture, the WLC handles the roaming process. The split-MAC architecture is discussed in Chapter 20.

  Because in a WLC environment the APs have established CAPWAP tunnels back to the WLC, the WLC maintains a database containing information on how to reach each wireless client. When a wireless client moves from one AP to another, the controller updates the client association from AP 1 to AP 2. Because both APs are bound to the same WLC, this is termed *intra-controller roaming*.

  When both APs are bound to the same controller, the roaming process is efficient. The process is efficient in the sense that it takes approximately 10 seconds, which is the time the WLC takes to switch the client entry from one AP to another. The wireless client does the roaming based on its own signal analysis, and it has no idea that the APs are connected with CAPWAP tunnels to a WLC.

  Along with the authentication communication between a WLC and a RADIUS server, cryptographic keys need to be generated and exchanged between a client and an AP or a WLC. Key exchanges during a roaming operation take a significant amount of time. Cisco WLCs have three techniques to help shorten some of the time and effort during roaming:

  - **Cisco Centralized Key Management (CCKM):** CCKM is a fast and secure roaming method created by Cisco. It uses a rekeying method to allow a client to roam from one AP to another without going through the controller. It reduces the time required for the client to authenticate with a new AP and derive a new session key during reassociation.
  - **Key caching:** With key caching, each wireless client maintains a list of keys used with previous associations with an AP and then presents them as it roams between them. The APs also maintain a database of keys issued to the client.
  - **802.11r:** This is the IEEE standard for fast roaming. It uses a process called Fast Transition (FT), where the initial handshake with the new AP is done before the wireless client roams to the target AP. The client

and the AP are basically doing a Pairwise Transient Key (PTK) calculation in advance. This PTK is applied to the client and AP after the client does the reassociation exchange with the new AP.

**Note**

The following three roaming scenarios cover WLAN roaming with multiple WLCs. When you have multiple WLCs, you can have the APs paired to them in a distributed fashion. Wireless clients can eventually roam from AP to AP from one WLC to another, depending on how the roaming target APs are paired to the WLCs. These three roaming scenarios continue from the previous two and deal with how you can use inter-controller roaming and mobility groups to coordinate roaming.

- **Layer 2 roaming between controllers:** Layer 2 roaming occurs when the WLAN interfaces of the WLCs are on the same IP subnet. When the wireless client associated with an AP joins a new controller, the new WLC exchanges mobility messages with the original WLC, and the client database entry is moved to the new WLC. If necessary, new security context and associations are established, and the client database entry is updated for the new AP. This process remains transparent from the wireless user's perspective as the two controllers coordinate the client's move between APs and controllers.
- **Layer 3 roaming between controllers:** Layer 3 roaming occurs when the WLAN interfaces of the WLCs are on different IP subnets. Layer 3 roaming is similar to Layer 2 roaming in the sense that the WLCs exchange mobility messages as the wireless client roams. However, instead of moving the wireless client database entry to the new WLC, the original WLC marks the wireless client with an "anchor" entry in its own client database. The database entry is copied to the new WLC client database and marked with a "foreign" entry in the new WLC. This process remains transparent from the wireless user's perspective, and the client maintains its original IP address.
- **Mobility groups:** A *mobility group* is a set of WLCs identified by the same mobility group name, and it defines the realm of seamless roaming for wireless clients. When you create a mobility group, you allow multiple WLCs in a network to dynamically share information and forward data traffic when inter-controller or inter-subnet roaming occurs. Controllers in the same mobility group can share the context and state of wireless client devices and their list of APs so that they do not consider

one another's APs as rogue devices. Using the learned information, the network can support inter-controller WLAN roaming and WLC redundancy.

Finally, you can use auto-anchor mobility (also called *guest tunneling*) to improve load balancing and enhance security for roaming wireless clients. When you use the auto-anchor mobility feature, you can specify a WLC or group of WLCs as the anchor points for wireless clients. Without this feature, a client that joins a WLAN is anchored to the first WLC. If the client then roams to a different subnet, the target WLC sets up a session for the client with the anchor WLC.

# Troubleshooting WLAN Configuration and Client Connectivity Issues

This final section of this chapter covers troubleshooting of wireless client connectivity issues for a single client and then for multiple clients from the WLC. It also covers troubleshooting of connectivity issues between a WLC and AP that may affect the connectivity of wireless clients.

The first step in troubleshooting wireless connectivity is to determine if the problem is isolated to a particular user, SSID, AP, or WLC. For example, suppose a single user is complaining about a connectivity issue. In that case, it may be more feasible to troubleshoot that user's interaction with a particular AP than to troubleshoot at the controller level. You might consider the following:

- The wireless client needs to be within the RF range of an AP that it can associate with.
- The AP that the wireless client is trying to connect to must not be misconfigured or malfunctioning.
- The wireless client must be able to successfully authenticate on the wireless network.
- The wireless client needs to be able to receive an IP address from a configured DHCP server.

If you look at these areas and still cannot resolve the issue, or if multiple clients are affected, you may need to start looking at other pieces of the wireless infrastructure, including WLCs, APs, and back-end authentication servers.

In the case of multiple wireless clients having issues in the same area, a good starting point would be to check on the AP. There may be a broken radio on

the AP that is preventing wireless clients from receiving a signal. This type of issue might require a physical check on the AP.

In addition to isolating the problem to a defective AP radio, you might need to look further up the stack. If an AP is completely non-operational, you might need to check the access layer switch that it is connected to, its switch port, and the PoE setting for that port, using the information gathered from Cisco Discovery Protocol (CDP). If the AP is receiving power but is non-operational or if you have multiple APs with issues, you might need to look even further up the stack. You might, for example, need to check that the AP has proper connectivity back to the WLC because the APs need to be properly joined to a WLC before clients can access the network.

From the WLC, you can also use an AP's built-in spectrum analyzer to monitor the wireless channels and identify sources of interference.

From the WLC, you can check the noise level on a channel and view an index of air quality on a channel under the Clean Air section. You can use the Clean Air section on a WLC to gather detailed information on interfering devices that may have been detected. You can also see air quality reports with AP, channels, and interferers information. A channel should have a high air quality value (measured from 0 to 100, with 100 being the best). The index of the air quality measures how nearby and interfering devices affect the quality of performance of a channel.

Figure 5.6 shows the Monitor section for a Cisco WLC.

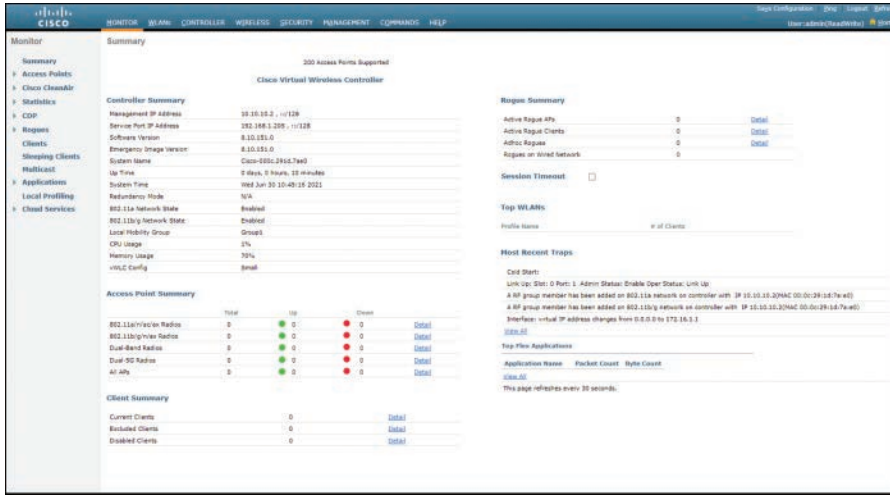



FIGURE 5.6 **Monitor Section for a Cisco WLC**

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. The decision to roam from one AP to another is initiated by which one of the following?
   - ❍ **A.** Wireless client
   - ❍ **B.** WLC
   - ❍ **C.** Current AP
   - ❍ **D.** Target AP

2. In terms of the time that it takes for a wireless client to roam, which of the following roaming methods is most efficient?
   - ❍ **A.** Layer 2 roaming
   - ❍ **B.** Layer 3 roaming
   - ❍ **C.** Intra-controller roaming
   - ❍ **D.** Mobility groups

3. When troubleshooting a wireless connectivity issue in a particular wireless coverage area, where should you start troubleshooting when multiple users are affected?
   - ❍ **A.** Wireless client
   - ❍ **B.** AP
   - ❍ **C.** Access switch
   - ❍ **D.** WLC

# Answers

1. **A** is correct. A wireless client would typically continuously scan the quality of the wireless connection, whether it is moving around or not, for better signal quality. If another AP provides a better RF signal, the wireless client reassociates to it.
2. **C** is correct. Because both APs are bound to the same controller in intra-controller roaming, the roaming process is more efficient.
3. **B** is correct. When multiple wireless clients are having issues in the same area, a good starting point would be to check on the AP.

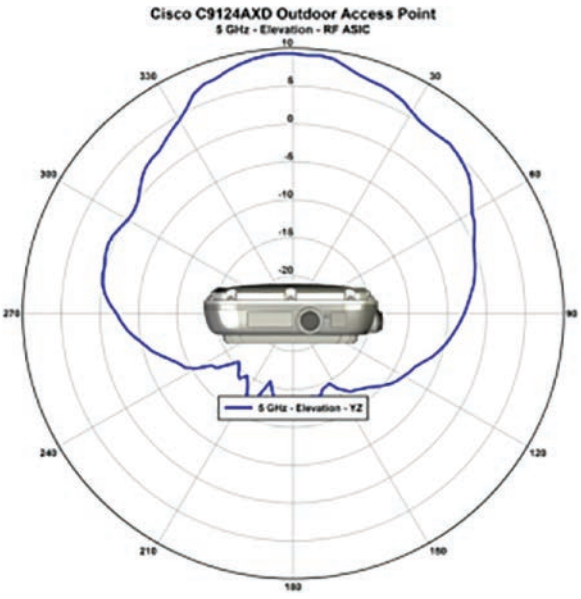
# Review Questions

1. What is the first step that a lightweight AP goes through after it boots up?
   - ❍ **A.** Configures itself with an IP address
   - ❍ **B.** Sets up a CAPWAP tunnel
   - ❍ **C.** Discovers a WLC
   - ❍ **D.** Downloads a configuration

2. True or false: When a pair of WLCs are set up in a high-availability SSO group, an AP needs to be joined to both controllers in the pair.
   - ❍ **A.** True
   - ❍ **B.** False

3. When troubleshooting a wireless connectivity issue on a single wireless client, where should you start troubleshooting?
   - ❍ **A.** Wireless client
   - ❍ **B.** AP
   - ❍ **C.** Access switch
   - ❍ **D.** WLC

4. Examine the following image.



What type of antenna would generate the signal pattern shown?

- ❍ **A.** Omnidirectional
- ❍ **B.** Directional
- ❍ **C.** Dipole
- ❍ **D.** Grid

# Answers to Review Questions

1. **A** is correct. The AP boots up and receives an IP address from a DHCP server if an IP address was not statically assigned previously.
2. **B** is correct. Deploying multiple WLCs in a high-availability configuration using stateful switchover (SSO) helps to alleviate the issue of an AP having to join a new WLC in the event of a WLC failure. The AP is joined to the active controller, and the configuration is synced to the standby controller.
3. **A** is correct. When a single wireless client is encountering connectivity issues, it may be most feasible to troubleshoot that wireless client interaction with a particular AP.
4. **B** is correct. The image shows the elevation pattern of a directional antenna.

# Further Reading

- ▶ *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers network device access control.

CHAPTER 6

# Device Access Control

**This chapter covers the following official ENCOR 350-401 exam objectives:**

- 5.1 Configure and verify devices access control
- 5.1.a Lines and password protection
- 5.1.b Authentication and authorization using AAA

This chapter is divided into two sections. The first section looks at the configuration and verification of network device access control with usernames and passwords. It also covers the configuration and verification of role-based access control (RBAC) using privilege levels. The second section covers authentication, authorization, and accounting (AAA). It looks at the configuration and verification of network device access control on Cisco IOS devices using TACACS+ and RADIUS.

**This chapter covers the following technology topics:**

- Cisco IOS CLI Session Overview
  - Protection of Access to Cisco IOS EXEC Modes
  - Secured Access with SSH
  - Privilege Levels and Role-Based Access Control (RBAC)
- Authentication, Authorization, and Accounting (AAA) Overview
  - TACACS+ Overview
  - RADIUS Overview
  - AAA Configuration for Network Devices

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. What are the first steps in securing user EXEC access to allow for secure network device access?
2. Which command option on remote CLI sessions is used to limit the session to use only a secure connection method?
3. What protocol does TACACS+ use for communication between a TACACS+ client (network device) and a TACACS+ server?
4. What are two of the high-level benefits of using a remote AAA server over local AAA services on each network device individually?

### Answers

1. Configure passwords for local and remote CLI sessions.
2. **transport input ssh**
3. TCP port 49
4. Scalability and standardized authentication methods using RADIUS and TACACS+

# Cisco IOS CLI Session Overview

Cisco IOS software provides several features that you can use to implement basic security for network devices' command-line sessions. These features include:

- Using different levels of authorization for CLI sessions to control access to commands that can modify the status of the networking device and for commands that are used to monitor the device
- Assigning passwords to CLI sessions
- Requiring users to log in to a networking device with a username
- Changing the privilege levels of commands to create new authorization levels for CLI sessions

You can establish IOS CLI sessions on Cisco IOS devices in two ways:

- **Local CLI sessions:** Local CLI sessions require direct access to the console port of the networking device. Local CLI sessions start in user EXEC mode. All of the tasks needed to configure and manage a networking device can be done using a local CLI session. The most common method for establishing a local CLI session is to connect a laptop to the console port of the networking device and then launch a terminal emulation application, like Putty, on the computer. The type of cable and connectors required and the settings for the terminal emulation application depend on the type of networking device that you are configuring. Some devices have an auxiliary (aux) port for remote administration through a dial-up modem. In most cases, this should be disabled with the **no exec** command under **line aux 0**.
- **Terminal lines and remote CLI sessions:** A remote CLI session is created between a host and a networking device by using a remote terminal access application, such as Telnet or SSH. Most of the tasks required to configure and manage a networking device can be done using a remote CLI session. The exceptions are tasks that interact directly with the console port (such as recovering from a corrupted operating system by uploading a new OS image over the console port) and interacting with the networking device when it is in ROMMON mode. SSH is a more secure alternative to Telnet. SSH provides encryption for the session traffic between the local management device and the networking device you are managing. Encrypting the session traffic with SSH prevents anyone who may have intercepted the traffic from decoding it.

With Cisco IOS networking devices, the word "lines" is used to refer to the software components that manage local and remote CLI sessions. You use the **line console 0** global configuration command to enter line configuration mode to configure options such as a password for the console port. Remote CLI sessions use lines that are referred to as vty lines. You use the **line vty** *line-number* [*ending-line-number*] global configuration command to enter line configuration mode to configure options such as a password for remote CLI sessions. Once you are in the line configuration mode, you can set the protocol you will be connecting over (for example, SSH).

Example 6.1 shows the console, auxiliary, and vty lines in the running configuration that are available on R1.

EXAMPLE 6.1 **Console, Auxiliary, and vty Lines in the Running Configuration**

```
R1#
R1# show running-config | section line
line con 0
line aux 0
line vty 0 4
R1#
```

Before we look at how to protect access to Cisco IOS EXEC modes, let's take a look at the five different types of passwords available in Cisco IOS:

- **Type 0 passwords:** Type 0 passwords are not encrypted and are stored in plaintext in the device configuration. The **enable password** command uses type 0 passwords. Type 0 passwords should not be used in a production environment.
- **Type 5 passwords:** Type 5 passwords use an MD5 hashing algorithm. These passwords are easily reversible with tools available on the Internet. The **enable secret** and **username** *username* **secret** commands use type 5 passwords.
- **Type 7 passwords:** Type 7 passwords uses the Vigenère cipher encryption algorithm, which is known to be weak. These passwords are easily reversible (in under 1 second) with tools available on the Internet. Type 7 password encryption is enabled with the **service password encryption** command.
- **Type 8 passwords:** Type 8 passwords use a Password-Based Key Derivation Function 2 (PBKDF2) with a SHA-256 hashed secret. Type 8 password security is considered good.
- **Type 9 passwords:** Type 9 passwords use the SCRYPT hashing algorithm. Type 9 passwords are considered the best passwords and should be used when supported.

Type 4 passwords were deprecated in IOS 15.3(3). The type 4 password hash was weaker than the type 5 (MD5) hash. Therefore, type 4 passwords should never be used. IOS 15.3(3) introduced support for type 8 and type 9 passwords, and these password types should always be used when supported.

# Protection of Access to Cisco IOS EXEC Modes

This section looks at the steps you can take to secure both user and privileged EXEC modes.

The first step in creating secure network device access is to protect the user EXEC mode by configuring passwords for local and remote CLI sessions. You start by entering line configuration mode by selecting the line number for the console port (for example, **line console 0**). Once you are in that mode, you use the **password** command to assign a password to **line console 0**. You use the **login** command at **line console 0** to enable password checking at login.

Next, let's look at configuring a password for remote CLI sessions. After a password is configured for remote CLI sessions, the IOS device prompts for a password the next time you establish a remote CLI session with that device. Cisco IOS networking devices require that a password be configured for remote CLI sessions. If you attempt to start a remote CLI session with a device that does not have a password configured for remote CLI sessions, you get a message indicating that a password is required and that the password is not set. The remote CLI session will be terminated by the remote host.

To configure a password for remote CLI sessions, you start by entering the line configuration mode and selecting the vty line (for example, **line vty 0 4**). When you are in that mode, you use the **password** command as you do for the console line. You use the **login** command at the vty line to enable password checking at login.

Example 6.2 shows how to assign a password to the console, auxiliary, and vty lines and verify it in the running configuration.

EXAMPLE 6.2 **Configuring and Verifying Line Passwords**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# line con 0
R1(config-line)# password Cisco123
R1(config-line)# login
R1(config-line)# line aux 0
R1(config-line)# password Cisco123
R1(config-line)# login
R1(config-line)# line vty 0 4
```

```
R1(config-line)# password Cisco123
R1(config-line)# login
R1(config-line)# end
R1#
R1# show running-config | section line
line con 0
 password Cisco123
 login
line aux 0
 password Cisco123
 login
line vty 0 4
 password Cisco123
 login
R1#
```

The previous section covers protection of access to both local and remote CLI sessions in user EXEC mode using line passwords. Now let's look at how to protect access to privileged EXEC mode. To add an additional layer of security, particularly for passwords that cross a network or that are stored with the configuration on a TFTP server, you can use the **enable secret** global configuration command.

Cisco recommends the use of the **enable secret** command over the **enable password** command because it uses an improved encryption algorithm. When you configure the **enable secret** command, it takes precedence over the **enable password** command. The two commands cannot be in effect simultaneously.

Let's look at the use of the **enable password** command to configure a password for privileged EXEC mode. The password you enter with the **enable password** command is stored as plaintext in the device's running configuration. You can encrypt the password for the **enable password** command in the configuration file of the networking device by using the **service password-encryption** command. However, the type 7 encryption level used by the **service password-encryption** command can be decrypted using tools available on the Internet, so it is not recommended for production deployments. The recommendation is to use the **enable secret** command because it provides strong encryption by hashing the password using type 5 passwords by default. However, on modern platforms, you can use type 8 or 9 passwords as well. You configure a password in privileged EXEC mode by using the command **enable secret** [**level** *level*] *unencrypted-password* | *encryption-type encrypted-password.* You can use the **show privilege** command to display the current level of privilege.

Example 6.3 shows the configuration and verification of protection of privileged EXEC mode using the **enable password** command. Note in the

verification that the password is stored in the running configuration in plaintext. This is because the default password, of type 0, was used. You can also set a type 7 password or set the EXEC level here. The command **service password-encryption** would make the password unreadable in the running configuration.

EXAMPLE 6.3 **Protecting Privileged EXEC with enable password**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# enable password ?
  0       Specifies an UNENCRYPTED password will follow
  7       Specifies a HIDDEN password will follow
  LINE    The UNENCRYPTED (cleartext) 'enable' password
  level   Set exec level password

R1(config)# enable password ExamCram123
 WARNING: Command has been added to the configuration using a type 0
password. However, type 0 passwords will soon be deprecated. Migrate
to a supported password type
R1(config)#
*Oct 28 23:00:00.922: %AAAA-4-CLI_DEPRECATED: WARNING: Command has
been added to the configuration using a type 0 password. However, type
0 passwords will soon be deprecated. Migrate to a supported password
type

R1(config)# do show run | include password
enable password ExamCram123
R1(config)#
R1(config)# service password-encryption
R1(config)# do show run | include password
enable password 7 106B11180834000A01557878
R1(config)# end
R1#
```

Example 6.4 shows the configuration and verification of protection of privileged EXEC mode using the **enable secret** command. This provides stronger encryption and is the recommended method to use. This example uses type 9 encryption. When using type 9, you need to type in the encrypted password or use the **algorithm-type** command to hash a plaintext **enable** secret. Note that the verification output shows the encrypted type 9 password.

EXAMPLE 6.4 **Protecting Privileged EXEC with enable secret**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
```

```
R1(config)# enable ?
  algorithm-type  Algorithm to use for hashing the plaintext 'enable'
secret
  password        Assign the privileged level password (MAX of 25
                  characters)
  secret          Assign the privileged level secret (MAX of 25
                  characters)

R1(config)# enable algorithm-type scrypt secret ?
  LINE   The UNENCRYPTED (cleartext) 'enable' secret
  level  Set exec level password

R1(config)# enable algorithm-type scrypt secret ExamCram123
R1(config)# do sho run | include secret
enable secret 9 $9$QlfhhreZrBM56f$VX4YG.yR/jHO/3gLFfTPqAw.
cdraNRDSKJoEOtCrC3Q
R1(config)# end
R1#
```

After you have protected access to user EXEC mode and privileged EXEC mode by configuring passwords for them, you can further increase the level of security on the device by creating usernames. You configure usernames to limit access to CLI sessions to a networking device to specific users. This is especially important if you are configuring a device to allow first-line technical support user access. These users typically would not need to run all commands available in privileged EXEC mode. For example, suppose you want technical support staff to be able to view the configuration on a device that will help them to troubleshoot network problems without being able to modify the configuration. In this case, you can create a username, configure it with privilege level 15, and configure it to run the **show running-config** command automatically. When a user logs in with the username, the running configuration will be displayed automatically.

There are three ways you can configure a username on a Cisco IOS device:

- Using the command **username** *username* **password** *password* configures a plaintext password (type 0).
- Using the command **username** *username* **secret** *password* provides type 5 encryption.
- Using the command **username** *username* **algorithm-type** [**md5** | **sha256** | **scrypt**] **secret** *password* provides type 5, type 8, or type 9 encryption, respectively.

The last option provides the highest level of security since it allows for the highest level of password encryption (type 8 or type 9). If the final option is not supported on a network device, then the second option should be used since it provides MD5 encryption. The first option should be avoided because it configures a plaintext password.

When you enable password authentication on a line by using the **password** command, you need to enable password checking. You do so by using the **login** command. This is what allows password use on the line. Once you have an alternate connection to the device, you can test the login. It is a good idea to have an alternate connection to a device if there is a problem logging in again using the line you made the changes on. The **login local** command allows for username/password pairs stored locally on the router to be used for the lines. By using the command **login local**, you can disable any password configured on lines.

To enable username and password authentication on a line, you need to do the following configuration:

- Create the user with the **username** command in global configuration mode, using one of the three options listed earlier in this section.
- Use the **login local** command in line configuration mode.

For remote CLI sessions, you can further protect the lines by using the **transport input** command. This command controls what protocols are allowed to access the vty lines. This can be configured with the command **transport input** {**all** | **none** |**telnet** | **ssh**}. The **all** option allows both Telnet and SSH access; **none** blocks Telnet and SSH; **telnet** allows only Telnet; and **ssh** allows only SSH access. Using **telnet ssh** allows both Telnet and SSH access. For the most secure access, the vty lines should be limited to SSH.

Example 6.5 shows the configuration and verification of usernames. The user **user1** is configured with a type 0 password, **admin1** is configured with a type 9 password, **tier1admin** is configured with a type 9 password (scrypt in this case), and **tier2admin** is configured with a type 8 password (sha256 in this case). The **login local** command is configured under the vty lines to tell it to use the router local user account database for authentication.

In this example, take note of the configured user accounts and the password types. **user1** with the type 0 password is shown in running configuration in plaintext. Privilege level 15 gives access to all commands, such as the **reload** command, and allows a user to make configuration changes on the device.

EXAMPLE 6.5 **Configuring Usernames and Passwords**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# username user1 password weakpassword
 WARNING: Command has been added to the configuration using a type 0
password. However, type 0 passwords will soon be deprecated. Migrate
to a supported password type
R1(config)# username admin1 privilege 15 secret admin1secret
R1(config)# username tier1admin algorithm-type scrypt secret
tier1adminsecret
R1(config)# username tier2admin algorithm-type sha256 secret
tier2adminsecret
R1(config)# line vty 0 4
R1(config-line)# login local
R1(config-line)# end
R1#
R1# show running-config | include username
username user1 password 0 weakpassword
username admin1 privilege 15 secret 9 $9$iVS2wE3FxvTvDv$6k.
NoCSCi2af4T8HpWeO1lBaTUnJzelT8S6xEETp7AI
username tier1admin secret 9 $9$bIFEJkC8eW9Xyf$vXBZD.8ZSiHTcjpNVfuMWwX
vveegKfHCfNXgLZUYA9w
username tier2admin secret 8 $8$PLF4/9DTLkfoTf$820AEmeaZA2mNh1oNJjAYk6
bYKSlLhUn9pULnifodyo
R1#
```

Example 6.6 shows how to establish a Telnet session from R2 to R1 by using username-based authentication with the **tier1admin** username and type 9 password created earlier. You can see here that you can successfully connect and authenticate by using the **tier1admin** account.

EXAMPLE 6.6 **Verifying Username-Based Authentication for vty Lines**

```
R2#
R2# telnet 100.1.1.1
Trying 100.1.1.1 ... Open


User Access Verification

Username: tier1admin
Password:

! Password entered is not displayed by the router
R1>

R1#
```

```
R1# show line
  Tty Typ     Tx/Rx  A Modem  Roty AccO AccI  Uses  Noise Overruns Int
*   0 CTY            -    -      -    -    -    0      0    0/0      -
    1 AUX 9600/9600  -    -      -    -    -    0      0    0/0      -
*  578 VTY           -    -      -    -    -    2      0    0/0      -
   579 VTY           -    -      -    -    -    0      0    0/0      -
   580 VTY           -    -      -    -    -    0      0    0/0      -
   581 VTY           -    -      -    -    -    0      0    0/0      -
   582 VTY           -    -      -    -    -    0      0    0/0      -

Line(s) not in async mode -or- with no hardware support:
2-577
! the * in the output of the showline command indicates that the first
vty (0) is in use
! vty 0 is mapped to vty 578 automatically

R1#
```

**ExamAlert**

For the ENCOR exam, it is important to know the differences between the two SSH versions as well as the high-level steps for SSH configuration on Cisco devices.

# Secured Access with SSH

SSH is a far more secure option than Telnet. Although Telnet is the most popular protocol used to access Cisco IOS devices, it is an insecure protocol. Its session packets are carried in plaintext, making it easy for someone to sniff and capture session information as it traverses the network. SSH provides encryption for session traffic between a device and a terminal access application. This prevents others from being able to intercept and decode the traffic.

SSH is available in two versions:

- **SSH Version 1 (SSHv1):** SSHv1 should be avoided because there are some flaws in its implementation, including its weak CRC-32 integrity check.
- **SSH Version 2 (SSHv2):** SSHv2 should be used when it is supported. The SSHv2 enhancement for RSA supports RSA-based public key authentication for a client and a network device. SSHv2 is not compatible with SSHv1.

Let us now take a look at the steps that are needed to set up a Cisco IOS device to run SSH:

1. Configure a hostname for the device, using the **hostname** *hostname* command.
2. Configure a domain name for the device, using the **ip domain-name** *domain-name* command.
3. Generate an RSA crypto key. Generating a key pair on the IOS device automatically enables SSH. When you generate an RSA key, you are prompted to enter a modulus length. A longer modulus length takes longer to generate, but it is more secure. You generate an RSA key with the **crypto key generate rsa** command.

Those three steps are mandatory. After you have taken those steps, you may need to set SSH to Version 2 because it is at SSHv1 by default on some platforms. You do this with the **ip ssh version 2** command. The other settings you can configure for the SSH service running on a device are the SSH timeout value and the authentication retries number. You do so with the command **ip ssh timeout** *seconds* **authentication-retries** *number*. Next, you set the transport input at the vty lines by using the **transport input ssh** command. Finally, also at the vty lines, you use the **login local** command to cause the local username and password on the router to be used for authentication.

For verification, you can use the **show ip ssh** command to view the version and configuration information for the SSH server. We can also use the **show ssh** command to show the status of the SSH server.

Example 6.7 demonstrates how to configure SSH, secure the vty lines to allow only SSH access, and verify connectivity from R2 to R1.

EXAMPLE 6.7 **Configuring and Verifying vty Access with SSH**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# username admin2 secret Cisco123
R1(config)# ip domain-name cisco.com
R1(config)# crypto key generate rsa
The name for the keys will be: R1.cisco.com
Choose the size of the key modulus in the range of 360 to 4096 for
your General Purpose Keys. Choosing a key modulus greater than 512 may
take a few minutes.
```

```
How many bits in the modulus [512]: 2048
% Generating 2048 bit RSA keys, keys will be non-exportable...
[OK] (elapsed time was 2 seconds)

R1(config)# ip ssh version 2
R1(config)# line vty 0 4
R1(config-line)# login local
R1(config-line)# transport input ssh
R1(config-line)# end
R1#

R2# ssh ?
  -c    Select encryption algorithm
  -l    Log in using this user name
  -m    Select HMAC algorithm
  -o    Specify options
  -p    Connect to this port
  -v    Specify SSH Protocol Version
  -vrf  Specify vrf name
  WORD  IP address or hostname of a remote system

R2# ssh -l admin2 -v 2 100.1.1.1


Password:
! Password entered is not displayed by the router

R1>
```

Finally, you can set a timeout for EXEC sessions that are left idle, which may pose a security risk. Under the line confirmation mode, you can use the **exec-timeout** *minutes seconds* command to set the timeout. The default setting is 10 minutes. Using **exec-timeout 0 0** and **no exec-timeout** disables the EXEC timeout. You should not use these commands this way in a production environment.

The **absolute-timeout** *minutes* command in the line configuration mode sets the interval for closing the EXEC session after a specified time has elapsed. This session is closed even if it is being used at the time of termination. You can use the **logout-warning** *seconds* command with the **absolute-timeout** command to notify users of an impending logout. By default, the user is given 20 seconds' notice before the session is terminated.

Example 6.8 shows how to configure EXEC and absolute timeouts and logout warning. For **line con 0**, a timeout value of 4 minutes is configured. For the vty lines, a value of 3 minutes and 30 seconds is configured. For the vty lines,

an absolute timeout of 10 minutes is configured, with a 120-second logout warning.

EXAMPLE 6.8 **Configuring EXEC and Absolute Timeouts**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# line con 0
R1(config-line)# exec-timeout 4 0
R1(config-line)# line vty 0 4
R1(config-line)# exec-timeout 3 30
!next we configure absolute timeout and logout warning
R1(config-line)# absolute-timeout 10
!logout warning is configured in seconds
R1(config-line)# logout-warning 120
R1(config-line)# end
R1#
```

# Privilege Levels and Role-Based Access Control (RBAC)

Now that we have examined the various ways of securing user and privileged EXEC modes, let's take a look at the use of privilege levels and RBAC. By default, Cisco IOS devices have three privilege levels:

- **Privilege level 0:** Privilege level 0 allows for the use of five commands: **enable**, **disable**, **help**, **logout**, and **exit**.
- **Privilege level 1:** Privilege level 1 is the user EXEC mode that you saw configured earlier in this chapter, in the section "Protection of Access to Cisco IOS EXEC Modes." In this mode, it is not possible to make configuration changes.
- **Privilege level 15:** Privilege level 15 is the privileged EXEC mode you saw configured earlier in this chapter, in Example 6.5. (It is also configured in the next example.) In this mode, all of the IOS CLI commands are available.

The commands that you can run in user EXEC mode at privilege level 1 are a subset of the commands that you can run in privileged EXEC mode at privilege 15. You can configure additional privilege levels from 2 through 14 to provide customized access control. For example, you might want to allow a group of

technical support staff to configure only a specific set of interface-level commands on interfaces while preventing device-wide configuration privileges. You could configure this in global configuration mode by using the command **privilege** *mode* **level** *level* [*command string*]. After you create that technical support user and assign this privilege, the user will be allowed to enter the interface and execute the commands specified in the command string. You can verify the configuration with the **show privilege** command.

Example 6.9 shows how to set up privileges to allow a network operation staff member to do basic manipulation of an interface. This example shows how to create the user **user1noc** with a type 9 password and privilege level 5 configured. In this particular case, a user with the **user1noc** username will be allowed to shut, unshut, and assign an IP address on the interface because these are the only commands this configuration allows in privilege level 5 in interface configuration mode. A user who tries to type a command that is not allowed (such as the **description** command) gets the message "Invalid input detected."

EXAMPLE 6.9 **Configuring and Verifying a Username and a Privilege Level**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# username user1noc privilege 5 algorithm-type scrypt secret
Cisco123
R1(config)# privilege exec level 5 configure terminal
R1(config)# privilege configure level 5 interface
R1(config)# privilege interface level 5 shutdown
R1(config)# privilege interface level 5 no shutdown
R1(config)# privilege interface level 5 ip address
R1(config)# end
R1#

R2# telnet 100.1.1.1
Trying 100.1.1.1 ... Open

User Access Verification

Username: user1noc
Password:

R1# show privilege
Current privilege level is 5
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
```

```
R1(config)# interface GigabitEthernet 0/0

!The options to configure on the interface are limited
R1(config-if)# ?
Interface configuration commands:
  default   Set a command to its defaults
  exit      Exit from interface configuration mode
  help      Description of the interactive help system
  ip        Interface Internet Protocol config commands
  no        Negate a command or set its defaults
  shutdown  Shutdown the selected interface

R1(config-if)# description test
                ^
% Invalid input detected at '^' marker.
R1(config-if)# end
R1#
```

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which of these commands can you use to force the vty lines to only allow remote connections via a protocol that supports encryption?
   - ❍ **A. transport input telnet**
   - ❍ **B. transport input ssh**
   - ❍ **C. crypto key generate rsa**
   - ❍ **D. ip ssh version 2**

2. What type of encryption does the **service password encryption** command provide?
   - ❍ **A.** Type 0
   - ❍ **B.** Type 5
   - ❍ **C.** Type 7
   - ❍ **D.** Type 9

3. True or false: SSH Version 1 implementation is compatible with SSH Version 2 implementation.
   - ❍ **A.** True
   - ❍ **B.** False

# Answers

1. **B** is correct. You can restrict the terminal line for SSH only by using the **transport input ssh** command in line configuration mode.
2. **C** is correct. Type 7 password encryption is enabled with the **service password encryption** command.
3. **B** is correct. SSHv2 is not compatible with SSHv1.

# Authentication, Authorization, and Accounting (AAA) Overview

Using line and local authentication as well as privilege levels works fine for controlling access on a small number of devices. However, this solution does not scale well as the number of devices grows. It becomes cumbersome and introduces the risk of inconsistent access control configurations across devices. To help simplify configuration and maintain consistency as the number of Cisco IOS devices grows, you can use an authentication, authorization, and accounting (AAA) solution.

There are many AAA protocol implementations, but this chapter focuses on the two most popular of them: RADIUS and TACACS+.

With AAA, network devices use a centralized RADIUS or TACACS+ server to authenticate users, authorize the commands users can run on a device, and provide accounting information. As a fallback mechanism, it is recommended that you still use local authentication in case the AAA server becomes unavailable at some point.

Let's briefly examine the AAA framework and how each part of it provides security functions:

- **Authentication:** Authentication provides identity verification before access to a network device is granted. It is the process of verifying the identity of the person or device accessing a network device, and it is based on the username and password combination provided by the entity trying to gain access.
- **Authorization:** Authorization provides access control. It is the process of assembling a set of attributes that describes what the user is authorized to perform. RADIUS and TACACS+ authorize users for specific rights by associating attribute/value (AV) pairs, which define the rights and the appropriate users.
- **Accounting:** Accounting provides a method for collecting information, logging the information locally on a network device, and sending the information to an AAA server for billing, auditing, and reporting. The accounting feature tracks and maintains a log of every management session used for access. You can use this information to generate reports for troubleshooting and auditing purposes.

Some of the high-level benefits of using a remote AAA server over local AAA services on each network device individually are highlighted next:

- Increased flexibility and control of access configuration
- Scalability
- Standardized authentication methods using RADIUS and TACACS+
- Ease of setup, since RADIUS and TACACS+ may have already been deployed across the enterprise
- More efficiency, since you can create user attributes once centrally and use them across multiple devices

Next, let's touch on the high points of TACACS+ and RADIUS before looking at their configuration.

# TACACS+ Overview

TACACS+ implementation provides for separate and modular authentication, authorization, and accounting facilities. It allows for a single access control server (referred to as the TACACS+ daemon) to provide authentication, authorization, and accounting to the network access server (NAS) independently. Typically, a client of a TACACS+ server is referred to as a NAS. A NAS may be a router, a switch, or an access point.

The TACACS+ protocol uses TCP port 49 for communication between the TACACS+ client (network device) and the TACACS+ server. A network administrator typically uses a workstation using Telnet, SSH, or the console to connect to a Cisco IOS device that needs to be managed. In this process, the TACACS+ client communicates with the TACACS+ server using the TACACS+ protocol. The TACACS+ protocol ensures confidentiality because all protocol exchanges between a TACACS+ client and a TACACS+ server are encrypted.

# RADIUS Overview

The Cisco implementation of RADIUS provides for a RADIUS client that runs on a Cisco IOS device to send an authentication request to a central RADIUS server that contains all user authentication and network service access information. RADIUS can be used with other AAA security protocols, such as local username lookup and TACACS+.

There are two implementations of RADIUS: Cisco's implementation and the industry-standard implementation. Cisco's implementation uses UDP port

1645 for authentication and authorization and UDP port 1646 for accounting. The industry-standard implementation uses UDP port 1812 for authentication and authorization and UDP port 1813 for accounting. The industry-standard implementation of the RADIUS protocol provides the distinction of working in a multi-vendor environment. Network devices from different vendors can connect to the same RADIUS server for AAA services. RADIUS can also be more convenient for AAA than TACACS+ since some organizations may already have it deployed.

As it relates to the privilege levels examined earlier in the chapter, TACACS+ and RADIUS can also be implemented when using AAA. For example, TACACS+ provides two ways to control the authorization of the network device commands on a per-user or per-group basis. One way is to assign privilege levels to commands and have the router verify with the TACACS+ server whether the user is authorized at the specified privilege level. Another way is to explicitly specify in the TACACS+ server, on a per-user or per-group basis, the allowed commands.

Cisco's TACACS+ and RADIUS implementations used to occur through the implementation of Cisco Secure Access Control Server (ACS), where RADIUS was used for network access control and TACACS+ was used for network devices access control. However, Cisco Identity Services Engine (ISE) is now the preferred implementation for AAA servers to support both TACACS+ and RADIUS protocols.

# AAA Configuration for Network Devices

In this section, you will see how both TACACS+ and RADIUS are configured from a Cisco IOS device. This section does not cover the configuration of a TACACS+ or RADIUS server because that is beyond the scope of this chapter.

There are two parts to configuring TACACS+ support: a TACACS+ server (for example, Cisco ISE) and a Cisco IOS device. At a high level, to configure a Cisco IOS device to support TACACS+, the following steps are involved:

1. Create a local user that will serve as the fallback if the TACACS+ server is not available or if you accidentally lock yourself out after enabling the AAA command. As highlighted previously, this is done with the command **username** *username* **privilege 15 algorithm-type {md5 | sha256 | scrypt} secret** *password*.
2. Enable the AAA function with the **aaa new-model** global configuration command.

3. Add a TACACS+ server.
4. Define the method lists for TACACS+ authentication by using the **aaa authentication** global configuration command.
5. Use the **line** and **interface** commands to apply the defined method lists to various interfaces.
6. If needed, use the **aaa authorization** global command to configure authorization for the device. Unlike with authentication, which can be configured per line or per interface, authorization is configured globally for an entire device.
7. If needed, use the **aaa accounting** command to enable accounting for TACACS+ connections.

Example 6.10 shows how to configure an IOS device with TACACS+ for device access control based on these steps. This example demonstrates basic authentication, authorization, and accounting configuration. Once the command **aaa-new model** is configured, there is no line authentication anymore on the vty lines as the default login method becomes AAA. The console port defaults to no authentication. If you were to disable this with the **no aaa new-model** command afterward, the login method would switch back to line authentication. However, you would not see **login local** under vty line; you would see just **login** (meaning just the line password will be checked, and not the local user database that is configured locally on the router).

EXAMPLE 6.10 **Configuring TACACS+**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
!Authentication setup
!First we create a fallback user account
R1(config)# username fallback privilege 15 algorithm-type scrypt
secret Cisco123
R1(config)# aaa new-model
R1(config)# tacacs server TACACSSERVER1
R1(config-server-tacacs)# address ipv4 100.1.1.2
R1(config-server-tacacs)# key Cisco123
R1(config-server-tacacs)# exit
R1(config)# aaa group server tacacs+ TACACSGROUP1
R1(config-sg-tacacs+)# server name TACACSSERVER1
R1(config)# aaa authentication login default group TACACSGROUP1 local
!the default method list automatically applies to all lines, except
the ones that have a named method list explicitly define or in other
words, it gets applied unless a more specific named method list is
defined.
```

```
!we can also specify on the vty lines the login authentication METH-
ODLIST1 command then tacacs+ TACACSGROUP1 will be used as the primary
authentication method and the local user database is set as the backup
R1(config)# line vty 0 4
R1(config-line)# login authentication methodlist1
!Authorization setup
!Next, for authorization we create a method list TACACSAUTH1
!if-authentication option allows a user who is authenticated to be
placed in EXEC mode
R1(config)# aaa authorization exec TACACSAUTH1 group TACACSGROUP1
local if-authenticated
R1(config)# aaa authorization commands 15 TACACSAUTH1 group TACACS-
GROUP1 local
!The config-commands command indicates that the server must return
permission to use any router configuration command
R1(config)# aaa authorization config-commands
R1(config)# aaa authorization console
!The TACACSAUTH1 method list is applied to the vty lines for both EXEC
and level 15 command access
R1(config)# line vty 0 4
R1(config-line)# authorization exec TACACSAUTH1
R1(config-line)# authorization commands 15 TACACSAUTH1
R1(config-line)# exit
R1(config)#
!Accounting setup
 !Next, for accounting we create a method list TACACSACC1
 !User EXEC sessions will be recorded as they start and stop, along
with user information
R1(config)# aaa accounting exec TACACSACC1 start-stop group
TACACSGROUP1
!commands that are entered while a user is in privilege level 15
(enable mode) will be recorded
R1(config)# aaa accounting commands 15 TACACSACC1 start-stop group
TACSRVGROUP1
!The TACACSACC1 method list is applied to the vty lines for EXEC and
level 15 commands
R1(config)# line vty 0 4
R1(config-line)# accounting exec TACACSACC1
R1(config-line)# accounting commands 15 TACACSACC1
R1(config-line)# end
R1#
```

The AAA server also needs to be configured with the AAA client information (that is, the hostname, IP address, and key), the login credentials for the users, and the commands the users are authorized to execute on the device.

At a high level, to configure a Cisco IOS device to support RADIUS, the following steps are involved:

1. Enable AAA with the **aaa new-model** global configuration command.
2. Define the RADIUS server and specify the IP address and key.
3. Add the RADIUS server to a server group.
4. Define method lists for RADIUS authentication by using the **aaa authentication login** *method-list* global configuration command.
5. Create a named method list and add a RADIUS server group as the primary and local database as backup by using the **aaa authentication login** command.
6. Use the **line** and **interface** commands to enable the defined method lists to be used. For example, Example 6.11 specifies the **login authentication** *method-list* command on the vty lines, and then the RADIUS server group will be used as the primary authentication method, and the local user database is set as the backup.

Example 6.11 shows the configuration of an IOS device with RADIUS for device access control based on these steps (which are nearly identical to the steps for TACACS+ configuration). This example demonstrates basic authentication configuration.

EXAMPLE 6.11 **Configuring RADIUS**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# aaa new-model
R1(config)# radius server RADIUSSERVER1
R1(config-radius-server)# address ipv4 100.1.1.2
R1(config-radius-server)# key Cisco123
R1(config-radius-server)# exit
R1(config)# aaa group server radius RADIUSGROUP1
R1(config-sg-radius)# server name RADIUSSERVER1
R1(config-sg-radius)# exit
R1(config)# aaa authentication login METHODLIST2 group RADIUSGROUP1
local
!the default method list automatically applies to all lines, except
the ones that have a named method list explicitly define or in other
words, it gets applied unless a more specific named method list is
defined.
!we can also specify on the vty lines the login authentication METHOD-
LIST2 command then RADIUSGROUP1 will be used as the primary authenti-
cation method and the local user database is set as the backup
```

```
R1(config-line)# line vty 0 4
R1(config-line)# login authentication METHODLIST2
R1(config-line)# end
R1#
```

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which of the following is not one of the benefits of AAA?
   - ❍ **A.** Increased flexibility and control of access configuration
   - ❍ **B.** Scalability
   - ❍ **C.** Standardized authentication methods using RADIUS and TACACS+
   - ❍ **D.** Complete removal of the need for local user creation on IOS devices

2. In the industry-standard implementation of the RADIUS protocol, which port is used for accounting?
   - ❍ **A.** UDP port 1645
   - ❍ **B.** UDP port 1646
   - ❍ **C.** UDP port 1812
   - ❍ **D.** UDP port 1813

3. Which command is entered to enable AAA on a Cisco IOS device?
   - ❍ **A. aaa authentication**
   - ❍ **B. aaa authorization**
   - ❍ **C. aaa new-model**
   - ❍ **D. aaa accounting**

## Answers

1. **D** is correct. As a fallback mechanism, it is recommended that you use local authentication in case the AAA server becomes unavailable at some point.
2. **D** is correct. The industry-standard implementation of RADIUS uses UDP port 1813 for accounting.
3. **C** is correct. When configuring both TACACS+ and RADIUS, you enable AAA functionality by using the **aaa new-model** global configuration command.

# Review Questions

1. In implementing the TACACS+ protocol, which port is used for communication between a network device and a TACACS+ server?
   - ❍ **A.** UDP port 1645
   - ❍ **B.** TCP port 49
   - ❍ **C.** TCP port 389
   - ❍ **D.** UDP port 1813

2. In TACACS+ implementation, which of the following can serve as network access servers?
   - ❍ **A.** Routers
   - ❍ **B.** Switches
   - ❍ **C.** Access points
   - ❍ **D.** All of the above

3. Which of the following commands is used for configuring a vty line to use the method list name **list1**?
   - ❍ **A. aaa authentication**
   - ❍ **B. aaa authorization**
   - ❍ **C. login authentication list1**
   - ❍ **D. aaa new-model**

4. To add a TACACS+ server in IOS 15.x, what command follows **tacacs server** *name* if the IP address is 10.10.10.10?
   - ❍ **A. aaa tacacs 10.10.10.10**
   - ❍ **B. server 10.10.10.10**
   - ❍ **C. address ipv4 10.10.10.10**
   - ❍ **D. aaa server 10.10.10.10**

# Answers to Review Questions

1. **B** is correct. The TACACS+ protocol uses TCP port 49 for communication between a TACACS+ client (network device) and a TACACS+ server.
2. **D** is correct. The clients of a TACACS+ server is referred to as a network access server (NAS). A NAS may be a router, a switch, or an access point.
3. **C** is correct. A method list enables logic authentication. To apply a custom list to a line, you use **login authentication** *custom-list name* in line configuration mode.

4. **C** is correct. To add a TACACS+ server in IOS 15.x, you need to specify the TACACS+ server name, specify the server IP address with the **address ipv4** *ip address* command (**address ipv4 10.10.10.10** in this case), and then specify the key string.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers infrastructure security.

CHAPTER 7

# Infrastructure Security

**This chapter covers the following official ENCOR 350-401 exam objectives:**

- 5.2 Configure and verify infrastructure security features
- 5.2.a ACLs
- 5.2.b CoPP

This chapter covers the configuration and verification of router security features. In particular, it looks at how you can use access control lists (ACLs) and control plane policing (CoPP). ACLs are versatile and can be used for managing various situations that involve controlling traffic on a router or switch. Apart from forwarding and blocking traffic on device interfaces or VLANs, ACLs can manage traffic to the logical construct of a router, the control plane. CoPP increases the security on a router by protecting the route processor (RP) from unnecessary or denial-of-service (DoS) traffic by prioritizing important control plane traffic.

**This chapter covers the following technology topics:**

- Access Control Lists (ACLs) Overview
  - Types of ACLs
  - Port ACLs (PACLs) and VLAN ACLs (VACLs)
- Control Plane Policing (CoPP)

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. What is one of the reasons you would use named access lists over numbered access lists?
2. What command is used to apply port access control lists (PACLs) to interfaces?
3. What are the main reasons you would implement the Cisco IOS control plane policing (CoPP) feature?
4. Which command is used to verify service policy implementation on the control plane for CoPP?

### Answers

1. Named access lists allow you to reorder statements in or add statements to an access list.
2. **ip access-group** *access-list* **in**
3. The Cisco IOS CoPP feature increases security on a router or switch by protecting the RP from unnecessary or denial-of-service (DoS) traffic and prioritizes important control plane and management traffic.
4. **show policy-map control-plane**

# Access Control Lists (ACLs) Overview

Access control lists (ACLs) can be used to perform packet filtering to control which packets move through a network device. An ACL is a sequential list of access control entries (ACEs) that permit or deny traffic on conditional matching statements. Processing starts at the top and proceeds downward until a matching pattern is identified (that is, from the lowest sequence to the highest sequence). Once a match is found, a permit or deny action is taken, and processing stops. At the end of each ACL is an implicit deny statement, which denies all traffic that did not match entries earlier in the ACL.

The following are some of the high-level benefits of using ACLs:

- **Blocking unwanted traffic:** Access lists can block incoming or outgoing traffic on an interface, thereby controlling access to a network based on

source and destination addresses, upper layer protocols, port numbers, differentiated services code point (DSCP), and so on.

- **Controlling access to vty lines:** Access lists on an inbound vty line can control who can access the vty line of a device. Access lists on an outbound vty line can control which destinations the lines from a device can reach.
- **Identifying or classify traffic for QoS:** Access lists provide congestion avoidance by setting IP precedence for weighted random early detection (WRED) and committed access rate (CAR). An access list can also provide congestion management for class-based weighted fair queueing (CBWFQ), priority queueing, and custom queueing.
- **Limiting debug command output:** Access lists can limit debugging output based on IP address or a protocol.
- **Providing bandwidth control:** Access lists can be used on a slow link to prevent excess traffic on a network.
- **Providing NAT control:** Access lists can control which addresses are translated using Network Address Translation (NAT).
- **Reducing the chance of DoS attack:** You can use access lists to limit the type of traffic permitted to the control plane of a device when configuring CoPP.
- **Restricting the content of routing updates:** Access lists can control routing updates sent, received, or redistributed in networks.

Here are some rules that apply to the implementation of access lists:

- Only one access list per interface, per protocol, and per direction is allowed.
- An access list must contain at least one **permit** statement, or all packets are denied due to the implicit deny entry.
- The order in which access list conditions or match criteria are configured is important. When deciding whether to forward or block a packet, Cisco devices check the packet against each criteria statement in the order in which these entries are created. After a match is found, no more entries are checked.
- If an access list is referenced by a name but does not exist, all packets pass. An interface or a command with an empty access list applied to it permits

all traffic into the network. (ACL types are covered in the following section.)

- Standard access lists and extended access lists cannot have the same name. (ACLs types are covered in the following section.)
- Inbound access lists process packets before the packets are routed to an outbound interface. Inbound access lists that have filtering criteria that deny packet access to a network are not subject to the overhead of routing lookups. Packets that are permitted access to a network based on the configured filtering criteria are processed for routing. When you configure a **permit** statement for inbound access lists, packets are processed after they are received, and when you configure a **deny** statement, packets are discarded.
- Outbound access lists process packets before they leave the device. Incoming packets are routed to the outbound interface and then processed by the outbound access list. When you configure a **permit** statement for outbound access lists, packets are sent to the output buffer, and when you configure a **deny** statement, packets are discarded.
- An access list can control traffic arriving at a device or leaving a device but not traffic originating at a device.

Before looking at the types of ACLs, let's briefly look at how to use wildcard masking for addresses in an access list. You use wildcard masking to indicate whether to check or ignore corresponding IP address bits when comparing the address bits in an access list entry to a packet being submitted to the access list. It is crucial to be careful when setting the wildcard mask, as you can specify one or more IP addresses for permit or deny checks.

With wildcard masking for IP address bits, you use the numbers 1 and 0 to specify how the router treats the corresponding IP address bits. A wildcard mask is sometimes referred to as an inverted mask because 1 and 0 mean the opposite of what they mean in a subnet mask:

- A wildcard mask bit 0 means check the corresponding bit value; these bit values must match.
- A wildcard mask bit 1 means ignore that corresponding bit value; these bit values need not match.

With a standard access list, if you do not supply a wildcard mask with a source or destination address in an access list statement, the router assumes an implicit wildcard mask of 0.0.0.0, meaning all values must match.

Unlike subnet masks, which require contiguous bits indicating network and subnet to be 1s, wildcard masks allow for non-contiguous bits in the mask.

> **ExamAlert**
>
> You should at least know the different ways to match with a wildcard mask for the ENCOR exam. You need to know how wildcard masks work and that they are powerful enough to match on odd and even IP addresses within a subnet.

Wildcard masks are quite powerful compared to subnet masks. You can use a wildcard mask to match on all odd or even numbers in a subnet. Say that you have a subnet where you want to permit IP addresses that have odd numbers in the fourth octet. Rather than writing an ACL with a lot of entries, you can use one statement. For example, say that you have a subnet 192.168.0.0/24 and only want IP addresses 192.168.0.1, 192.168.0.3, 192.168.0.5, and so on to be allowed. If you are to write these out in binary, all of the addresses with the odd number in the fourth octet have the last bit as 1. To permit all IP addresses with the fourth octet as odd, this last bit should remain as is. Thus, you can use a wildcard mask that cares only about the last bit and about nothing in the fourth octet. The statement can be written as **access-list 20 permit 192.168.0.1 0.0.0.254**.

You can use a similar technique if you are trying to match on all even numbers in a subnet. For example, if you are trying to match on 192.168.0.0, 192.168.0.2, 192.168.0.4, and so on, you can write these out in binary and see that the even number in the fourth octet has the last bit as 0. Similarly, to match on the odd number, you can use a wildcard mask that cares about the last bit and does not care about any bit in the fourth octet. The statement can be written as **access-list 20 deny 192.168.0.0 0.0.0.254**.

Table 7.1 shows examples of IP addresses and subnet masks from an access list, along with the corresponding addresses that are considered matches.

TABLE 7.1 **Address, Wildcard Mask, and Match Results**

| Address | Wildcard Mask | Match Result |
|---|---|---|
| 0.0.0.0 | 255.255.255.255 | All addresses match the access list conditions |
| 172.16.0.0/16 | 0.0.255.255 | Network 172.16.0.0 |
| 172.16.1.2/16 | 0.0.0.0 | Only host 172.16.1.2 matches |
| 172.16.2.0 | 0.0.0.7 | Only subnet 172.16.2.0/29 matches |
| 172.16.2.15 | 0.0.0.3 | Only subnet 172.16.2.15/30 matches |

ExamAlert

For the ENCOR exam, it would be helpful to know the different types of ACLs, their number ranges, and the features they provide.

# Types of ACLs

Various kinds of ACLs can be used for packet filtering. This section gives you a closer look at numbered standard ACLs, numbered extended ACLs, and named ACLs. Broadly, all access lists are either standard or extended. If you intend to filter on a source address, a standard access list will suffice. For filtering on anything other than source, you need to use an extended access list. Named access lists can be either standard or extended access lists. You will also learn about port ACLs (PACLs) and VLAN ACLs (VACLs) in the following section.

Generally, numbered access lists are specified as either standard or extended, based on the number in the **access-list** command. Let's now look at standard, extended, and named access lists in more depth.

## Standard ACLs

Standard ACLs are numbered 1 to 99 or 1300 to 1999; the range of standard IP access lists was initially only 1 to 99 but was expanded later. Standard IP access lists check only source addresses of packets (with two exceptions). Since standard access lists check source addresses, they are very efficient at blocking traffic closer to a destination. These are the exceptions when the address in a standard access list is not a source address:

- On outbound vty access lists, when someone is trying to use Telnet, the address in the access list entry is used as a destination address rather than as a source address.
- When filtering routes, you are filtering the network being advertised to you rather than a source address. For example, when filtering routes in route distribution between OSPF and another routing protocol, you can match the network to be filtered based on a standard ACL or a prefix list.

These are the steps to create and apply a numbered standard access list to an interface:

1. Define the ACL by using the command **access-list** *access-list-number* {*deny* | *permit*} {*source* [*source-wildcard*] | **any**} [**log**]. A standard access list needs

to be numbered from 1 to 99 or from 1300 to 1999. You can optionally use the keyword **any** as a substitute for *source/source-wildcard* to specify the source and source wildcard of 0.0.0.0 255.255.255.255.

2. Repeat step 1 until you have specified the sources on which you want to base the access list. Remember that all sources not specifically permitted are denied by an implicit **deny** statement at the end of the access list.
3. If needed, display the contents of all current IP access lists with the **show ip access-list** command.
4. Apply the access list to an interface by using the command **ip access-group** *access-list number* {**in** | **out**}.

## Extended ACLs

Extended ACLs are numbered 100 to 199 or 2000 to 2699. Like the standard access list range, the extended access list range was eventually expanded from its original range of 100 to 199. Extended access lists are suitable for blocking traffic anywhere. Extended access lists check source and destination addresses and other IP packet data, such as protocols, TCP or UDP port numbers, type of service (ToS), precedence, TCP flags, and IP options. Extended access lists can also provide capabilities that standard access lists cannot, such as the following:

- Filtering IP options
- Filtering TCP flags
- Filtering noninitial fragments of packets
- Providing time-based entries

These are the steps to create and apply a numbered extended access list to an interface:

1. Define the ACL by using the command **access-list** *access-list-number* {**deny** | **permit**} **protocol** {*source* [*source-wildcard*] | **any**} {*destination* [*destination-wildcard*] | **any**} [**precedence** *precedence*] [**tos** *tos*] [**established**] [**log** | **log-input**] [**time-range** *time-range-name*] [**fragments**]. An extended access list needs to be numbered from 100 to 199 or from 2000 to 2699. The **log-input** keyword is used to include the input interface, source MAC address, or virtual circuit in the logging output.

2. If desired, add a remark by using the command **access-list** *access-list-number* **remark** *remark*. The **remark** command adds a user-friendly comment about an access list entry.
3. Repeat steps 1 and 2 until you have specified the fields and values on which you want to base the access list.
4. Apply your numbered extended access list to an interface by using the command **ip access-group** *access-list number* {**in** | **out**}.

Example 7.1 demonstrates how a numbered extended access list is created and applied to an interface. This access list blocks Telnet and ICMP traffic and allows all other traffic. This example includes an optional remark that describes the entry.

EXAMPLE 7.1 **Creating and Applying a Numbered Extended Access List**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# access-list 100 deny tcp any any eq 23
R1(config)# access-list 100 deny icmp any any
R1(config)# access-list 100 permit ip any any
R1(config)# access-list 100 remark block_telnet_icmp
R1(config)# interface GigabitEthernet 0/0
R1(config-if)# ip access-group 100 in
R1(config-if)# end
R1#
```

# Named ACLs

Named ACLs can be specified as either standard or extended, with the **standard** and **extended** keywords in the **ip access-list** command. Every access list must be identified by either a name or a number. Named access lists are more convenient than numbered access lists because you can specify a meaningful name that is easier to remember and associate with a task. You can reorder statements or add statements to a named access list. Named access lists support the following features that are not supported by numbered access lists:

- IP options filtering
- Noncontiguous ports
- TCP flag filtering
- Deletion of entries with the **no permit** or **no deny** command

These are the steps you follow to create and apply a named extended access list to an interface:

1. Define the ACL name with the command **ip access-list extended** *name*. This defines an extended IP access list using a name and enters extended named access list configuration mode.
2. In the access list line configuration mode, define the ACL by using the command {**permit** | **deny}** *protocol source* [*source-wildcard*] *destination* [*destination-wildcard*] [**option** *option-name*] [**precedence** *precedence*] [**tos** *tos*] [**established**] [**log** | **log-input**] [**time-range** *time-range-name*] [**fragments**].
3. Repeat steps 1 and 2 until you have specified the fields and values on which you want to base the access list.
4. Apply the access list to an interface by using the command **ip access-group** *access-list name* {**in** | **out**}.

Example 7.2 shows how named standard and extended access lists are created and applied to an interface.

EXAMPLE 7.2 **Creating and Applying Named Standard and Extended Access Lists**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# ip access-list standard NAMED_STANDARD_ACL_EXAMPLE
R1(config-std-nacl)# deny 10.10.20.0 0.0.0.255
R1(config-std-nacl)# deny host host 172.16.0.1
R1(config-std-nacl)# deny host 172.16.0.1
R1(config-std-nacl)# permit any
R1(config-std-nacl)# exit
R1(config)# interface GigabitEthernet 0/1
R1(config-if)# ip access-group NAMED_STANDARD_ACL_EXAMPLE in
R1(config-if)# end
R1#

R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# ip access-list extended NAMED_EXTENDED_ACL_EXAMPLE
R1(config-ext-nacl)# deny tcp any any eq 23
R1(config-ext-nacl)# deny icmp any any
R1(config-ext-nacl)# permit ip any any
R1(config-ext-nacl)# exit
```

```
R1(config)# interface GigabitEthernet 0/2
R1(config-if)# ip access-group NAMED_EXTENDED_ACL_EXAMPLE in
R1(config-if)# end
R1#
R1# show access-lists
Standard IP access list STANDARD_ACL_EXAMPLE
    20 deny   172.16.0.1
    10 deny   10.10.20.0, wildcard bits 0.0.0.255
    30 permit any
Extended IP access list 100
    10 permit tcp any any eq 22
    20 deny tcp any any eq telnet
    30 deny icmp any any
    40 permit ip any any
Extended IP access list NAMED_EXTENDED_ACL_EXAMPLE
    10 deny tcp any any eq telnet
    20 deny icmp any any
    30 permit ip any any
R1#
```

Cisco IOS software can provide logging messages about packets permitted or denied by a single standard or extended IP access list entry. The way this works is any packet that matches the entry causes an informational logging message about the packet to be sent to the console. The level of messages logged to the console is controlled by the **logging console** global configuration command (covered in Chapter 6, "Device Access Control").

The first packet that triggers the access list entry causes an immediate logging message, and subsequent packets are collected over 5-minute intervals and displayed or logged. The logging message includes the following:

- Access list number
- Whether the packet was permitted or denied
- Source IP address of the packet
- Number of packets from that source permitted or denied in the prior 5-minute interval

You can use the **ip access-list log-update** command to set the number of packets that, when matched on an access list (and permitted or denied), cause the system to generate a log message. You can adjust this if you want to receive log messages more frequently than at 5-minute intervals.

# Port ACLs (PACLs) and VLAN ACLs (VACLs)

Cisco Catalyst switches allow you to apply access lists to Layer 2 ports and to VLANs. The access lists applied to Layer 2 ports are called *port access control lists (PACLs)*, and the access lists applied to VLANs are known as *VLAN access control lists (VACLs)*.

## PACLs

A PACL provides the ability to perform access control on specific Layer 2 ports. A Layer 2 port can be a Catalyst switch physical port or trunk port that belongs to a VLAN. PACLs are applied only on ingress traffic. The PACL feature is supported only in hardware; that is, PACLs are not applied to any packets routed in software. You apply an IPv4 PACL to an interface by using the **ip access-group** *access-list* **in** command.

The IOS CLI syntax for creating a PACL is identical to the syntax for creating a Cisco IOS ACL. An instance of an ACL that is mapped to a Layer 2 port is called a PACL. An instance of an ACL mapped to a Layer 3 interface is called a Cisco IOS ACL. The same ACL can be mapped to both a Layer 2 port and a Layer 3 interface. However, the PACL feature supports MAC ACLs and IPv4 ACLs. The PACL feature does not support ACLs for IPv6, Address Resolution Protocol (ARP), or Multiprotocol Label Switching (MPLS) traffic.

Example 7.3 shows an extended access list being applied to a switch port. In this case, the extended ACL **PACL_EXAMPLE** is being used to deny Telnet and ICMP traffic into GigabitEthernet 0/1 on SW1.

EXAMPLE 7.3 **Extended Access List Applied on Switch Port (PACL)**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# ip access-list extended PACL_EXAMPLE
SW1(config-ext-nacl)# deny tcp any any eq 23
SW1(config-ext-nacl)# deny icmp any any
SW1(config-ext-nacl)# permit ip any any
SW1(config-ext-nacl)# exit
SW1(config)# interface GigabitEthernet 0/1
SW1(config-if)# ip access-group PACL_EXAMPLE in
SW1(config-if)# end
SW1#
```

# VACLs

VACLs can provide access control for all packets bridged within a VLAN or routed into or out of a VLAN. Unlike regular Cisco IOS ACLs that are applied on routed packets only, VACLs apply to all packets and can be applied to any VLAN. VACLs are processed in the ACL TCAM hardware, and VACLs ignore any Cisco IOS ACL fields that are not supported in hardware. You can configure VACLs for IP and MAC-layer traffic.

As you saw earlier, PACLs are implemented like regular Cisco IOS ACLs. However, the implementation of a VACL is slightly different. The following steps highlight how to filter traffic that is bridged within a VLAN or traffic that is routed into or out of a VLAN:

1. Define the VLAN access map by using the command **vlan access-map** *name sequence*. A VLAN access map consists of one or more VLAN access map sequences, where each VLAN access map sequence consists of one match and one action statement.
2. Configure the match statement by using the command **match** {**ip address** {*acl-number* | *acl-name*} | **mac address** *acl-name*}. The match statement supports standard, extended, or named IPv4 ACLs as well as named MAC address ACLs for the matching criteria.
3. Configure the action statement by using the command **action forward** | **drop** [**log**]. The action statement specifies the action that is taken when a match occurs (for example, forward traffic, drop traffic). You can use the **log** keyword to log dropped traffic.
4. Apply the VACL by using the command **vlan filter** *vlan-access-map-name vlan-list*. The option *vlan-list* can be a single VLAN, a range of VLANs (such as **2-10**), or a comma-separated list of multiple VLANs (such as **2,4-6,8**).

Example 7.4 demonstrates how to create and apply a VACL.

EXAMPLE 7.4 **Creating and Applying a VACL**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# ip access-list extended TELNET
SW1(config-ext-nacl)# permit tcp any any eq 23
SW1(config-ext-nacl)# exit
```

```
SW1(config)# ip access-list extended ICMP
SW1(config-ext-nacl)# permit icmp any any
SW1(config-ext-nacl)# exit
SW1(config)# ip access-list extended IP_TRAFFIC
SW1(config-ext-nacl)# permit ip any any
SW1(config-ext-nacl)# exit

SW1(config)# vlan access-map VACL_50 5
SW1(config-access-map)# match ip address TELNET
SW1(config-access-map)# action drop log
SW1(config-access-map)# exit

SW1(config)# vlan access-map VACL_50 10
SW1(config-access-map)# match ip address ICMP
SW1(config-access-map)# action drop log
SW1(config-access-map)# exit

SW1(config)# vlan access-map VACL_50 15
SW1(config-access-map)# match ip address IP_TRAFFIC
SW1(config-access-map)# action forward
SW1(config-access-map)# exit

SW1(config)# vlan filter VACL_50 vlan-list 50
SW1(config)# exit
SW1#
SW1# show vlan filter
VLAN Map VACL_50 is filtering VLANs:
  50
SW1#
```

Now let's now look at the order of operations when it comes to PACLs, VACLs, and Cisco IOS ACLs. With an incoming packet on a physical port, a PACL is applied first. If the PACL permits the packet, the VACL on the ingress VLAN is then applied. If the packet is Layer 3 forwarded and is allowed by the VACL, it is filtered by the Cisco IOS ACL on the same VLAN. The same process happens in reverse in the egress direction.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. How many access lists per protocol and per direction are allowed on an interface?
   - ❍ **A.** One
   - ❍ **B.** Two
   - ❍ **C.** Four
   - ❍ **D.** Unlimited

2. What command do you use to apply a VACL?
   - ❍ **A. ip access-group** *access-list name* {**in** | **out**}
   - ❍ **B. ip access-class** *access-list name* {**in** | **out**}
   - ❍ **C. ip access-list** *access-list name* {**in** | **out**}
   - ❍ **D. vlan filter** *vlan-access-map-name vlan-list*

3. Which of the following can a PACL be applied to? (Choose two.)
   - ❍ **A.** Layer 2 port
   - ❍ **B.** Layer 3 port
   - ❍ **C.** Trunk
   - ❍ **D.** VLAN

## Answers

1. **A** is correct. Only one access list per interface, per protocol, and per direction is allowed.
2. **D** is correct. The command **vlan filter** *vlan-access-map-name vlan-list* is used to apply a VLAN ACL to a single VLAN, a range of VLANs, or a comma-separated list of multiple VLANs.
3. **A** and **C** are correct. A PACL can be applied to the Layer 2 port of a Catalyst switch, including a physical port or trunk port that belongs to a VLAN.

# Control Plane Policing (CoPP)

The traffic that is managed by a device's route processor (RP) is divided into three broad functional planes: the data plane, the management plane, and the control plane.

Before we get into protecting the control plane with control plane policing (CoPP), let's do a brief refresher on the control plane.

Network device-generated or device-received packets that are used for the creation and operation of a network are referred to as *control plane packets*. The vast majority of packets are handled in the data plane of a router. However, the RP must handle certain packets, such as those related to routing updates and network management. From the perspective of a network device, control plane packets always are handled by the CPU in the network device RP. And because the RP is critical to network operations, service disruptions to it or to the control and management planes can cause network outages.

The following are classified as control plane traffic:

- Routing protocol traffic
- Packets destined to the local IP address of the router
- Simple Network Management Protocol (SNMP) packets
- Interactive access protocol traffic, such as Secure Shell (SSH) and Telnet, traffic
- Traffic related to protocols such as Internet Control Message Protocol (ICMP) or IP options that might also require handling by the device CPU
- Layer 2 protocol packets such as bridge protocol data unit (BPDU) and Cisco Discovery Protocol (CDP) packets

CoPP is a Cisco IOS-wide feature that is designed to allow users to manage the flow of traffic handled by the RP of a network device. It is designed to prevent unnecessary traffic from overwhelming the RP that, if left unabated, could affect device performance by exhausting resources on the device. *RP resource exhaustion*, in this case, refers to all resources associated with the punt path and RPs, such as Cisco IOS process memory and buffers and ingress packet queues.

More than just the control plane packets can be punted and affect the RP and system resources. Management plane traffic, as well as certain data plane exceptions, IP packets, and some service plane packets may also require the use of RP

resources. Even so, it is common practice to identify the resources associated with the punt path and RP as the control plane.

In summary, the Cisco IOS CoPP feature increases security on a router or switch by protecting the RP from unnecessary or denial-of-service (DoS) traffic and prioritizes important control plane and management traffic.

The following steps are involved in implementing CoPP:

1. After identifying the network traffic, create ACLs for matching in a class map. (ACLs are covered earlier in this chapter.)
2. Create class maps to use the created ACLs to match known protocols. The **class-map** command is used to define a traffic class, which contains three major elements: a name, one or a series of **match** commands (to specify various criteria for classifying packets), and an instruction on how to evaluate the **match** command(s). Packets are checked to see whether they match the criteria specified in the **match** commands. If a packet matches the specified criteria, that packet is considered a member of the class and is treated according to the QoS specifications set in the service policy. Packets that fail to meet any of the matching criteria are classified as members of the default class.
3. After classifying the traffic, create policy maps to enforce policy actions for the identified traffic. The **policy-map** command is used to associate a traffic class, defined by the **class-map** command, with one or more QoS policies. The result of this association is a *service policy*, which contains three elements: a name, a traffic class (specified with the **class** command), and the QoS policies. The purpose of a service policy is to associate a traffic class with one or more QoS policies. Classes included in policy maps are processed from the top down. When a packet is found to match a class, no further processing is performed. That is, a packet can only belong to a single class, and it is the first one to which a match occurs. When a packet does not match any of the defined classes, it is automatically placed in the class **class-default**. The default class is always applied, whether it is explicitly configured or not.
4. Use the **service-policy** command to attach the service policy, as specified with the **policy-map** command, to an interface. In the case of CoPP, this is the control plane interface. Because the elements of the service policy can be applied to packets entering—or, in some versions of CoPP, leaving—the interface, you are required to specify whether the service policy characteristics should be applied to incoming or outgoing packets.

In particular, you apply the CoPP policy map to the control plane by using the **service-policy {input | output}** *policy-name* command.

5. For verification, use the **show policy-map control-plane input** command.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. What is the name of the CoPP construct that ties together predefined ACLs?
   - ❍ **A.** ACL
   - ❍ **B.** Class map
   - ❍ **C.** Policy map
   - ❍ **D.** Service map

2. What command is used to apply a service policy to the control plane to implement CoPP?
   - ❍ **A. class-map**
   - ❍ **B. policy-map**
   - ❍ **C. service-policy**
   - ❍ **D. ip access-group**

3. True or false: The CoPP feature increases security on a router or switch by protecting the RP from unnecessary or denial-of-service (DoS) traffic.
   - ❍ **A.** True
   - ❍ **A.** False

## Answers

1. **B** is correct. Class maps use created ACLs to match known protocols, addresses, IP precedence, DSCP values, CoS, and so on.
2. **C** is correct. The **service-policy {input | output}** *policy-name* command is used to attach a service policy to the control plane.
3. **A** is correct. CoPP protects the RP from unnecessary or denial-of-service (DoS) traffic and gives priority to the important control plane and management traffic.

# Review Questions

1. Which types of ACLs are applied in the Layer 2 switch environment? (Choose two.)
   - ❍ **A.** Standard ACLs
   - ❍ **B.** Extended ACLs
   - ❍ **C.** PACLs
   - ❍ **D.** VACLs

2. What happens when a matching ACE is found in an ACL?
   - ❍ **A.** Action is taken, and processing is stopped on the remaining ACE.
   - ❍ **B.** Processing continues to the next ACE.
   - ❍ **C.** Regardless of matching statements, processing needs to go through all ACEs.
   - ❍ **D.** Processing continues through other ACEs when there is a **permit** statement.

3. A VACL VLAN list can reference all except which of the following?
   - ❍ **A.** A single VLAN
   - ❍ **B.** A range of VLANs
   - ❍ **C.** A comma-separated list of multiple VLANs
   - ❍ **D.** Layer 2 ports

4. True or false: CoPP is applied device-wide on a Cisco IOS device.
   - ❍ **A.** True
   - ❍ **B.** False

# Answers to Review Questions

1. **C** and **D** are correct. Access lists that are applied to Layer 2 ports are called port access control lists (PACLs), and access lists that are applied to VLANs are known as VLAN access control lists (VACLs).
2. **A** is correct. Once a matching ACE is found, a permit or deny action is taken, and processing stops.
3. **D** is correct. A VACL VLAN list can be a single VLAN, a range of VLANs, or a comma-separated list of multiple VLANs. It cannot reference Layer 2 ports.
4. **A** is correct. CoPP is a Cisco IOS-wide feature designed to be applied to the control plane to manage the flow of traffic handled by the RP of a device.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*
- *CCNP and CCIE Enterprise Core & CCNP Advanced Routing Portable Command Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers the securing of REST APIs.

*This page intentionally left blank*

CHAPTER 8

# Securing REST APIs

**This chapter covers the following official ENCOR 350-401 exam objective:**

- 5.3 Describe REST API Security

This chapter covers the Security section of the ENCOR 350-401 exam related to REST API security. This chapter looks at the basics of application programming interfaces (APIs) and representational state transfer (REST) APIs. Then it looks at how to secure the REST API communication channel from prying eyes and how to verify that data is from a trusted source.

**This chapter covers the following technology topic:**

- Securing REST APIs

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. Which API is responsible for communication from the SDN controllers to all the services that run over a network?
2. What type of identity certificate does ISE provide to the Cisco DNA Center to secure the communication?

### Answers

1. Northbound API
2. X509 v3

# REST API Security

Before getting into the security for REST APIs, let's review application programming interfaces (APIs) and representational state transfer (REST) APIs.

An API is programming code that generally allows two software programs to communicate. You can use APIs for configuring and monitoring network components. In a software-defined networking (SDN) architecture, there are three stacked layers. Let's look at them briefly, as understanding them will help you get a better grasp when we examine the northbound and southbound APIs shortly.

The three stacked layers of the SDN architecture are:

- **Data or forwarding plane:** The data or forwarding plane contains all the network components (both physical and virtual) involved in moving customer traffic. It simply executes the controller's rules and does not contain the intelligence of the SDN architecture.
- **Control plane:** The control plane is the core layer of the SDN platform. It includes the controllers, and its main job is to facilitate the creation and destruction of network paths.
- **Application plane:** The application plane enables SDN applications to communicate network requirements to the controller. Its responsibilities include management and reporting functions (including monitoring and security).

ExamAlert

Make sure you understand the two common APIs, Northbound and Southbound, for the ENCOR exam.

There are two common APIs between these layers:

- **Northbound API:** The northbound API is implemented by the controller of the SDN architecture and is responsible for communication from the SDN controllers to all the services that run over the network. An example of northbound API communication is using the Cisco DNA Center graphical user interface (GUI) to manage the network controller. A network engineer would log in to the controller to manage the network. The information that is passed between the controller and the web browser in this example leverages the northbound REST-based API.
- **Southbound API:** The southbound API is responsible for communication between the SDN controller and the network devices. If a network engineer needs to make a configuration change through the GUI on the controller's management interface, that would be an example of using the southbound API. The network change being pushed to the routers, switches, and wireless access points uses the southbound API. The following are a few examples of a few southbound APIs:
  - **OpenFlow:** OpenFlow, which was developed by the Open Networking Foundation (ONF), was one of the early southbound interfaces. With OpenFlow, the configuration of the network devices is done using NETCONF.
  - **OpFlex:** OpFlex is an open standard that allows policies to be applied across physical and virtual network devices in a multi-vendor environment. With OpFlex, policies are defined within a logical, centralized repository in the controller. The OpFlex protocol is used to communicate and enforce those policies with distributed elements in the network devices.
  - **NETCONF:** NETCONF uses a simple remote-procedure call (RPC) mechanism to facilitate communication between a client (for example, a centralized management platform script or application) and a server (for example, a network device). NETCONF is IETF standardized, and its messages are encoded using Extensible Markup Language (XML).

- **RESTCONF:** RESTCONF uses structured data (XML or JSON) and YANG to provide REST-like APIs, enabling you to programmatically access different network devices.

A REST API is a web-based API that expects some data as input (typically XML). Usually, the client sends data via an HTTP POST, HTTP GET, or even HTTP DELETE verb to a well-known URL. The data is in the format of plain XML, JSON, or text, with no particular envelope or wrappers. The processing happens on the server side. The server then sends back the data, typically in XML, JSON, or plaintext format.

**ExamAlert**

Before you take the ENCOR exam, ensure that you have a complete understanding of how to make REST API communication secured, especially using HTTPS to ensure that communication is encrypted using SSL/TLS.

When using automation, including APIs, you must consider the security of the communications between the controller and applications or network devices. There are two parts to making REST API communications secure:

- Securing the communication channel from prying eyes by using HTTPS
- Verifying that the data is from a trusted source

Most current APIs have the ability to protect data in motion. For example, let us examine the security of Cisco DNA Center using a REST API to push a configuration to the Cisco Identity Services Engine (ISE). In this setup, Cisco ISE acts as an HTTPS server, and the communication from Cisco DNA Center and Cisco ISE will be encrypted using Secure Sockets Layer (SSL) or Transport Layer Security (TLS). Using HTTPS ensures that communication is encrypted using SSL/TLS. It protects the API credentials and transmitted data.

Security best practices suggest that the communication between the software and the controller should be encrypted to ensure that someone with bad intentions who captures the data flow cannot easily view the data.

HTTPS also provides additional security so that Cisco ISE, acting as the server, provides an X509 v3 identity certificate to secure the communication. In this case, Cisco DNA Center must verify the identity certificate by validating the signature attached to the certificate before a secure channel is established for communication. This can be achieved by installing the trusted root certificate of the root CA that signed the certificate or by installing the

identify certificate itself if you are using a self-signed certificate. This secure communication ensures that Cisco DNA Center is pushing configuration to the trusted device.

In this case, it is also important to have Cisco ISE verify the identity of DNA Center before it applies the configuration that it receives. The REST API HTTPS requests are authenticated using HTTP basic authentication. Authentication is performed for every request using the basic authentication header added to every API request. In this example, for this to work, the administrative user must be created in Cisco ISE with permissions that enable access to the REST API interface.

Other built-in methods are used for securing a REST API. For example, DNA Center uses basic authentication to pass a username and password to the DNA Center Token API for users' authentication. This API allows for authentication to the DNA Center controller for additional API calls.

Similarly, you can use Postman (an API platform for building and using APIs) to pass credentials to the DNA Center controller via the Token API. When you successfully authenticate to the DNA Center controller, you receive a token that contains a string, and future API calls need a new token. You can think of it as a hash generated from the login credentials supplied. Using this method is secure as the token received is valid only for the current authenticated session to the controller. If another user accesses the controller, another unique token will be generated for that user session, based on the credentials supplied.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. True or false: An API generally allows two software programs to communicate with each other.
   - ❍ **A.** True
   - ❍ **B.** False

2. Which API is responsible for the communication between SDN controllers and network devices?
   - ❍ **A.** Northbound
   - ❍ **B.** Southbound
   - ❍ **C.** XML
   - ❍ **D.** JSON

# Answers

1. **A** is correct. An API is programming code that generally allows for two software programs to communicate with each other.
2. **B** is correct. The southbound API is responsible for communication between a controller of the SDN controllers and network devices.

# Review Questions

1. True or false: At a minimum, you should ensure that communication between a controller and software is encrypted with TLS.
   - ❍ **A.** True
   - ❍ **B.** False

2. True or false: When multiple users access a DNA Center controller, the same token is used.
   - ❍ **A.** True
   - ❍ **B.** False

# Answers to Review Questions

1. **A** is correct. Security best practices suggest that you protect the communication between the software and the controller by using HTTPS to ensure that communication is encrypted using SSL/TLS. This protects the API credentials and the transmitted data.
2. **B** is correct. When multiple users access a DNA Center controller, multiple unique tokens are generated for the different user sessions, based on the credentials supplied.

# Further Reading

- ▶ *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*
- ▶ *Network Programmability and Automation Fundamentals*

# What's Next?

If you want more practice on this chapter's exam objective before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers wireless security.

*This page intentionally left blank*

CHAPTER 9

# Wireless Security

**This chapter covers the following official ENCOR 350-401 exam objectives:**

- 5.4 Configure and verify wireless security features
- 5.4.a EAP
- 5.4.b WebAuth
- 5.4.c PSK

This chapter covers the configuration and verification of several wireless authentication methods commonly used in the enterprise wireless network environment. First, this chapter looks at Open Authentication, which authenticates wireless users without credentials. Next, it looks at authenticating wireless users with static pre-shared keys. With this option, clients can use a fixed text string that is common across wireless clients and access points (APs) to access a wireless network. Next, this chapter examines how to authenticate wireless clients using Extensible Authentication Protocol (EAP) connected to an external authentication server or using a local user database. Finally, the chapter wraps up by looking at authenticating wireless users by using WebAuth. These users are prompted for credentials or prompted to acknowledge a warning or an alert statement on a web page.

**This chapter covers the following technology topics:**

- Wireless Authentication Overview
  - Open Authentication
  - Pre-Shared Key (PSK) Authentication
  - Extensible Authentication Protocol (EAP) Authentication
  - WebAuth

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. Why should you avoid using Open Authentication by itself in an enterprise wireless deployment?
2. What are the two authentication modes available to you when using a version of WPA?
3. Where is the supplicant located when using 802.1X to authenticate wireless clients?
4. With WebAuth, which type of Layer 3 security authenticates wireless users against a local database?

### Answers

1. With Open Authentication, no authentication is performed by the wireless client before associating with an AP. This is a major security risk. However, Open Authentication can be used in collaboration with WebAuth so that, after the wireless client is associated automatically with the AP, the wireless user can be prompted on a web page to acknowledge an acceptable use policy before continuing.
2. Pre-Shared Key (Personal mode) and 802.1X (Enterprise mode)
3. On the wireless access client
4. Local web authentication with an internal database

# Wireless Authentication Overview

Before a wireless device can communicate over a network, it must first be associated with an access point (AP). This process starts by initially discovering the wireless LAN (WLAN) and connecting to it. During this initial process, the transmitted frames can reach any devices that are within range. It is crucial that the WLANs be secure so that only users who are authenticated and authorized are granted access to network resources. Once you properly secure a WLAN as well, you should be able to prevent malicious users from accessing network resources over the wireless network.

This same association process is also used for guest WLAN access. It is common in modern enterprise networks to have a guest service set identifier (SSID) deployed for visitors-only connectivity. This way, visitors can only

access nonconfidential, or public, resources. All other clients that are unknown or not welcomed should not be permitted to be associated with an AP at all. This chapter explores the wireless client's authentication process, which happens before client devices are allowed to access resources on a WLAN.

Four wireless authentication methods are covered in the following sections. The first three, Open Authentication, PSK, and EAP, provide Layer 2 security; the fourth, WebAuth, is considered Layer 3 security and is triggered by a client opening a web session (either HTTP or HTTPS).

# Open Authentication

**ExamAlert**

For the ENCOR exam, you need to have a complete understanding of why you should not use Open Authentication. You should also understand why you may want to use it in some cases and how to secure it (for example, by using VLAN segmentation and firewall rules) when you do.

With Open Authentication, no authentication is performed by the wireless client before associating with an AP. This allows for any device to be connected to that particular wireless network. Open Authentication and WEP were the only options available for authentication in the original 802.11 standards.

Generally, network administrators should avoid using Open Authentication because it does not verify a wireless device that is trying to associate with an AP. The only requirement is that the wireless client must use an 802.11 request before getting associated with the AP; no credentials are required. Open Authentication can be used in collaboration with WebAuth so that, after the wireless client is associated automatically with the AP, the wireless user can be prompted on a web page to acknowledge an acceptable use policy before continuing. However, the main takeaway about using Open Authentication is that the wireless client is associated automatically with the AP.

Typically, you should avoid using Open Authentication for obvious security reasons. However, it does provide simplicity, allowing you to have a client on the network without any complex configuration. For example, you might use Open Authentication when you are trying to get a guest onto a wireless network where ease of connectivity is paramount and access control is not required. Ideally, if you are going to use open authentication, you should configure VLAN segmentation and firewall rules in the back-end network to prevent unauthorized users with malicious intent from gaining access to corporate network resources.

There is not a whole lot that needs to be done in terms of configuration to set up a wireless network for Open Authentication. On a WLC, you need to create a WLAN and map it to a VLAN. Then you need to set up the SSID string, set the controller interface, and enable the network.

Figure 9.1 shows a WLAN name Open_WLAN, with SSID Open_WLAN and status set to Enabled and connected to a custom interface called Open on a Cisco WLC.

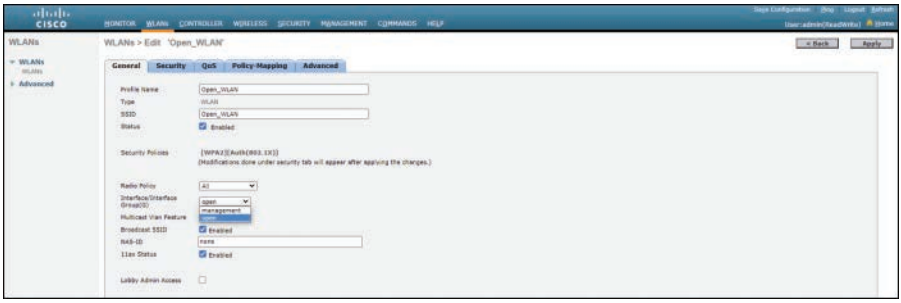



FIGURE 9.1 **Open Authentication Configuration**

Figure 9.2 shows the security configuration for WLAN Open_WLAN. In this case, the WLAN named Open_WLAN was predefined and mapped to a VLAN in the interface in Figure 9.1. You use the Security tab to set the wireless security and user authentication parameters. From the Layer 2 Security dropdown, you select None for Open Authentication.

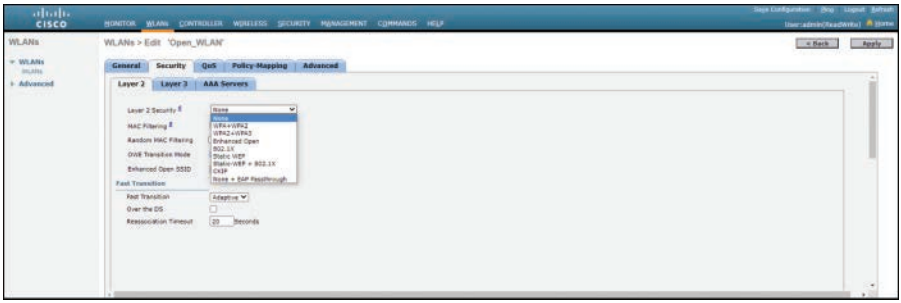



FIGURE 9.2 **Open Authentication Security Configuration**

You can verify this configuration by looking at the General tab and ensure that the security policy is shown as None, which indicates that you are using Open Authentication (see Figure 9.3).

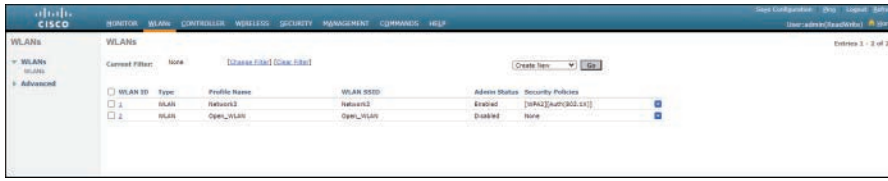


FIGURE 9.3 **Open Authentication Security Verification**

# Pre-Shared Key (PSK) Authentication

A pre-shared key (PSK) method of authentication allows anyone who has a key to access the network. Along with Wired Equivalent Privacy (WEP), which has been deprecated, you can use one of three Wi-Fi Protected Access (WPA) versions to secure data as it is being sent across a wireless network. In addition to authentication, each WPA version also specifies encryption and data integrity methods to protect data.

Depending on the scale and complexity of the wireless deployment, you can use two authentication modes when using WPA: pre-shared key or 802.1X. These are also known as Personal and Enterprise modes, respectively. A pre-shared key is a key string shared between a wireless client and an AP with which it is trying to associate. The key string is not passed over the air. Instead, the wireless client and the AP go through a four-way handshake process. The pre-shared key is used to generate encryption key material that is then exchanged between the wireless client and the AP. Once this process is complete, the two can securely transfer data over the air.

Before delving into the configuration of pre-shared key authentication, let's look briefly at wireless authentication types as well as the wireless encryption methods available. This recap should help you better understand PSK authentication.

The following methods are available to secure wireless connections:

- **Wired Equivalent Privacy (WEP):** WEP, which has been deprecated, was part of the original 802.11 standards. It used a 40- to 128-bit key that was a combination of a key and an initialization vector. WEP is still available to support some legacy devices, but it should be avoided as it can be easily decoded and has been deemed insecure.
- **Wi-Fi Protected Access (WPA):** WPA was created to address some of the problems with WEP. It included a new type of key called the Temporal Key Integrity Protocol (TKIP). TKIP develops a new, unique encryption key for each wireless frame, thus providing a more secure

connection. Nevertheless, WPA also should be avoided because TKIP is susceptible to wireless attacks.

- **WPA2 Personal:** WPA2 Personal uses the more secure Advanced Encryption Standard (AES) and pre-shared key for authentication. It is backward compatible with TKIP for interoperability.
- **WPA2 Enterprise:** WPA2 Enterprise utilizes user-level authentication along with 802.1X standards with AES encryption.
- **WPA3 Personal:** WPA3 Personal uses simultaneous authentication of equals (SAE) to build on WPA2 PSK to allow users to authenticate with a passphrase only. SAE provides resistance to offline dictionary attacks, where an attacker may try to determine the network password by guessing.
- **WPA3 Enterprise:** WPA3 Enterprise provides protection for a network transmitting sensitive data by offering 192-bit cryptographic strength. This is considered the most secure wireless authentication method.

Let us now take a look at the wireless encryption methods that are available today:

- **Temporal Key Integrity Protocol (TKIP):** TKIP is an encryption method used by WPA. It makes use of WEP but encrypts the Layer 2 payload and uses a message-integrity check to ensure that a packet has not been altered.
- **Advanced Encryption Standard (AES):** AES is an encryption method used by WPA2. AES uses counter mode with the Cipher Block Chaining Message Authentication Code Protocol (CCMP) encryption protocol, with a stronger algorithm for message integrity and confidentiality. CCMP allows the destination host to determine whether encrypted and non-encrypted bits have been altered in transmission.
- **Galois/Counter Mode (GCM):** GCM uses the same cryptographic engine as AES but embeds it in a more efficient framework, making it more secure than CCMP. GCM is used in WPA3.

You set WPA2 and WPA3 Personal configuration in the Layer 2 tab under the Security tab of the WLC. If you do not already have a WLAN, you need to create one, edit it, go to Security > Layer 2, and select the Layer 2 Security mode and enter the pre-shared key. Figure 9.4 shows the configuration of PSK

authentication on a WLAN named Employees. The version of WPA has been set for WPA2+WPA3 and Personal mode.

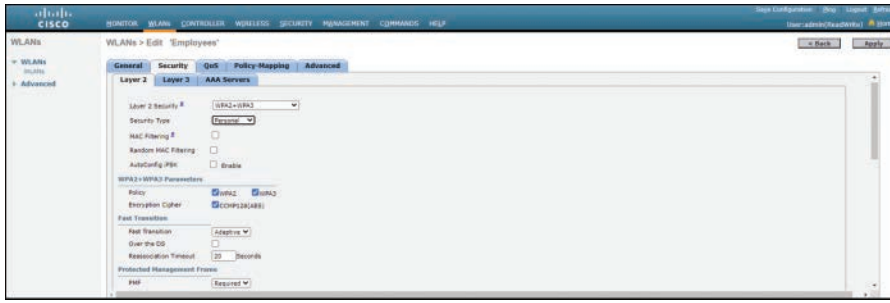


FIGURE 9.4 **WPA3 Personal Security Configuration**

Figure 9.5 shows the PSK setup in the Authentication Key Management section.

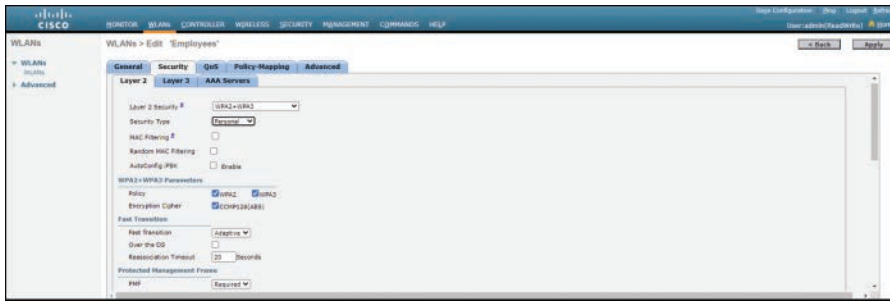


FIGURE 9.5 **Authentication Key Management for PSK**

Figure 9.6 shows the screen where you verify that the Employee WLAN has been set to PSK and WPA3 Personal.

FIGURE 9.6 **WPA3 Personal Verification**

# Extensible Authentication Protocol (EAP) Authentication

**ExamAlert**

For the ENCOR exam, you need to have a complete understanding of EAP and the components that are needed in an 802.1X setup, as well as their function.

Extensible Authentication Protocol (EAP) works with 802.1X to provide a flexible and scalable authentication framework for wireless users. EAP is extensible in the sense that it does not involve only one authentication method. Instead, it outlines a set of functions and authentication methods for authenticating wireless users.

EAP integrates with 802.1X by limiting access to a wireless network until the client can be authenticated. Wireless clients may be able to associate with an AP but unable to move data over the air until the client is authenticated. With 802.1X, the client uses Open Authentication to associate with an AP; then, the authentication happens on an authentication server (usually an external RADIUS server). 802.1X is covered further in Chapter 11, "Network Access Control."

The following components are needed in an 802.1X setup to securely authenticate wireless users:

- **Client (supplicant):** The supplicant is the client device that is requesting access to the network.
- **Authenticator:** The authenticator is the network device that is providing access to the network. An AP forwards the supplicant's message to the WLC, which is the authenticator.
- **Authentication server:** The authentication server is the device that accepts the user or client credentials and denies or permits access to the network based on policies and a user database. In most cases, this is a RADIUS server.

To implement EAP with 802.1X, you should use the highest WPA Enterprise mode that is supported between the WLC, AP, and wireless clients. The WLC can use either the local user database on the WLC or an external RADIUS server located on the network. Let's now take a look at the setup of this on a WLC. This section does not look at the configuration of the actual RADIUS server but only at how to define the RADIUS server(s) on the WLC.

To begin the configuration, you need to define one or more RADIUS servers on the WLC by navigating to Security > AAA > RADIUS > Authentication. You also need to configure other options, such as the RADIUS server priority, port, shared secret to communicate with the RADIUS server, and server status. You also need to check the **Enable** box next to Network Users to authenticate wireless users. You can add multiple RADIUS servers, and the WLC will use them in sequential order.

Figure 9.7 shows the configuration to define a RADIUS server on a WLC for WPA3 Enterprise authentication.

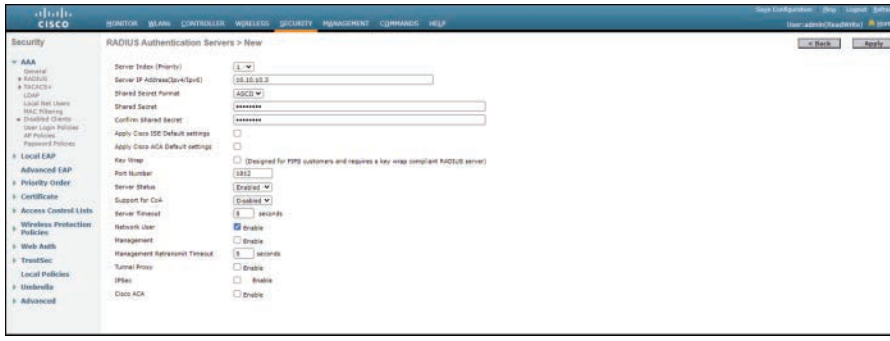



FIGURE 9.7 **RADIUS Server Configuration for WPA3 Enterprise Authentication**

Next, you need to create a WLAN and edit it on the Security tab by selecting a WPA type from the Layer 2 Security dropdown. Ideally, if the WLC, AP, and wireless clients support it, you should select WPA2+WPA3. To force the use of a WPA Enterprise mode, you have to select Enterprise from the Security Type dropdown list. Then you select the WPA policy and encryption to use. Figure 9.8 shows the configuration for WPA3 Enterprise.

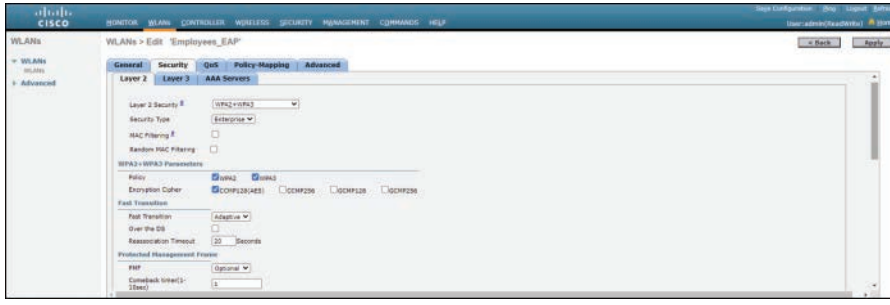



FIGURE 9.8 **WPA3 Enterprise Security Configuration**

Finally, under Authentication Key Management, you check the checkbox next to 802.1X (see Figure 9.9) to force the use of 802.1X.


FIGURE 9.9 **Authentication Key Management for 802.1X**

Let's now look at setting up EAP authentication with the WLC local user database. This is known as Local EAP. There are several reasons you might use Local EAP. For instance, maybe you have a small number of wireless users, or your environment is not large enough to justify RADIUS servers.

To configure Local EAP, you first need to enable it on the WLC. This is done by navigating to Security > Local EAP > Profiles and creating a new profile. Once this is complete, you should see the profile and the authentication method that it supports. Next, you create a WLAN and specify the use of Local EAP.

Before you can use Local EAP, you need to remove the RADIUS servers configured on the WLC. If both RADIUS and Local EAP are configured, the WLC will attempt to use RADIUS first, and if that fails, it will fall back to using Local EAP. To start using EAP, you need to check the Enable checkbox for Local EAP Authentication and select the EAP profile. You can navigate to this configuration page by going to WLAN > Security > AAA Servers > Local EAP Authentication.

The final step in setting up Local EAP is to define the users' accounts on the WLC. Because Local EAP uses the local interface user database on the WLC instead of external authentication such as RADIUS, you need to manually define all the users who require access to the WLAN on the WLC. You create users by navigating to Security > AAA > Local Net Users. On this page, you mainly need to create a username and a password and select the WLAN profile. The process of creating a Local EAP profile and defining the local net users is straightforward, so we skip over those steps. However, we do take a look at where to configure the WLAN to use Local EAP. Figure 9.10 shows how to enable Local EAP for a WLAN by using the authentication server built in to the WLC. On the WLAN Security tab, you enable EAP and select the EAP profile that was predefined. In this case, you are using the predefined EAP Profile LocalEAP1, which has EAP enabled.


FIGURE 9.10 **Enabling Local EAP on a WLAN**

# WebAuth

WebAuth presents wireless users with a web interface. A wireless user can enter credentials, read information about the organization, acknowledge an acceptable use policy, and so on before connecting. Or the web interface may simply present an acceptable use policy that a user must accept before connecting, without entering credentials; this is called *web passthrough mode*.

WebAuth is a Layer 3 security feature that accepts user credentials before granting access to a wireless network. It works like this:

1. The client sends a DNS query for a web server, and the WLC spoofs the DNS response and provides its own virtual IP address.
2. The client opens a web (HTTP or HTTPS) connection to that address.
3. The WLC redirects to itself (via Local Web Authentication) or to an external authentication server (via Central Web Authentication), which then prompts the wireless users to enter credentials.

We will look more closely at these steps shortly.

Before delving further into the WebAuth process, let's briefly look at the types of Layer 3 security available for accessing wireless networks:

- **Web passthrough:** Web passthrough is a slight variation of regular web authentication in which the user is not prompted to enter credentials. Typically, you would use some form of warning, alert, or acceptable use policy, but you don't have this with web passthrough.
- **Local web authentication with internal database:** Local web authentication validates a wireless user's credentials against the internal local database on the WLC.

- **Local web authentication with external database:** Local web authentication with an external database basically connects to a RADIUS or LDAP server.
- **Local Web Authentication with external redirect:** Local Web Authentication with external redirect means the login page for web authentication is stored on an external server.
- **Central Web Authentication with Cisco ISE:** Central Web Authentication means the authentication happens on Cisco ISE. A web portal in ISE provides the login page to the wireless user. After the user is verified in ISE, the client device is provisioned.

Let's look at the step-by-step process that occurs when a wireless user tries to connect to a wireless network configured for web authentication:

1. The wireless user opens a web browser and enters a URL, such as http://www.pearson.com. The client then sends out a DNS request for this URL to get the IP address for the destination. The WLC passes the DNS request to the DNS server, and the DNS server responds with a DNS reply, which contains the IP address of the destination www.pearson.com. This is then forwarded to the wireless client.
2. The client tries to open a TCP connection with the destination IP address. It sends out a TCP SYN packet destined to the IP address www.pearson.com.
3. The WLC, which has rules configured for the client and hence can act as a proxy for www.pearson.com, sends back a TCP SYN-ACK packet to the client with the source as the IP address of www.pearson.com. The client sends back a TCP ACK packet to complete the three-way TCP handshake, and the TCP connection is fully established.
4. The client sends an HTTP GET packet destined to www.pearson.com. The WLC intercepts this packet and sends it for redirection handling. The HTTP application gateway prepares an HTML body and sends it back as the reply to the HTTP GET requested by the client. This HTML makes the client go to the default web page URL of the WLC's virtual IP address, such as http://10.10.10.1/login.html.
5. The client closes the TCP connection with the IP address (for example, www.pearson.com).

6. The client wants to go to http://10.10.10.1/login.html and tries to open a TCP connection with the virtual IP address of the WLC. It sends a TCP SYN packet for 10.10.10.1 to the WLC.
7. The WLC responds with a TCP SYN-ACK, and the client sends back a TCP ACK to the WLC to complete the handshake.
8. The client sends an HTTP GET for /login.html destined to 10.10.10.1 to request the login page.
9. This request is allowed up to the web server of the WLC, and the server responds with the default login page. The client receives the login page on the browser window, where the wireless user can enter credentials.

For WebAuth configuration, you need to create a WLAN on a WLC and map it to a VLAN. Then you need to set up the SSID string, set the controller interface, and enable the network. On the Security tab and then the Layer 2 tab, you set the security to None. This forces the WLAN to use WebAuth. The Layer 3 tab is where you set up the different Layer 3 security types highlighted earlier. Figure 9.11 shows the web authentication configuration on the Guest WLAN.


FIGURE 9.11 **Web Authentication Configuration**

Finally, under Security > Web Auth > Web Login Page, you set the web content that will be displayed to wireless users by setting a headline and the message that you would like to be presented. This can be an acceptable use policy asking the wireless user to click on the Accept button before using the wireless network.

Figure 9.12 shows the configuration of the WebAuth login page.

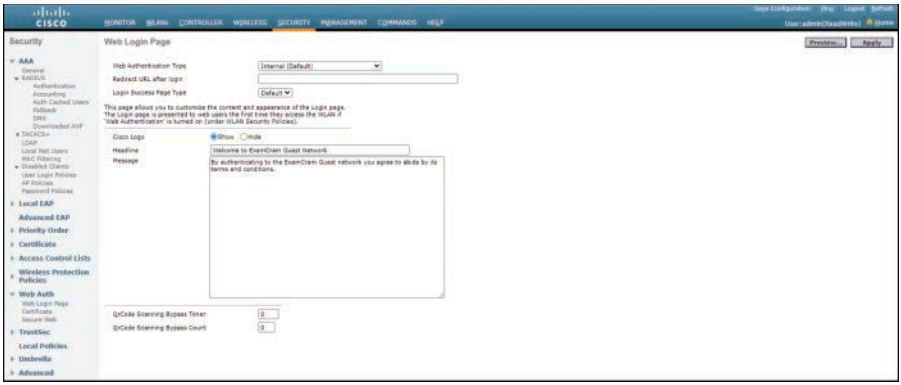



FIGURE 9.12 **Configuring WebAuth Login Page**

Figure 9.13 shows the WebAuth login page that will be presented to users of the Guest WLAN.


FIGURE 9.13 **Web Login Page Preview**

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. In Open Authentication, what is required before a wireless client can associate with an AP?

   ❍ **A.** PSK

   ❍ **B.** RADIUS

   ❍ **C.** 802.1x

   ❍ **D.** None of the above

2. With wireless client authentication, what two devices share the PSK?
   - ❍ **A.** WLC and AP
   - ❍ **B.** Wireless client and WLC
   - ❍ **C.** Wireless client and AP
   - ❍ **D.** WLC and RADIUS server

3. Which of the following wireless authentication methods is considered the most secure?
   - ❍ **A.** WPA2 Personal
   - ❍ **B.** WPA2 Enterprise
   - ❍ **C.** WPA3 Personal
   - ❍ **D.** WPA3 Enterprise

4. True or false: When using WebAuth, passthrough mode does not require wireless users to enter credentials on a web page.
   - ❍ **A.** True
   - ❍ **B.** False

# Answers

1. **D** is correct. With Open Authentication, the wireless client is automatically associated with the AP, and no further input is required from the wireless client.
2. **C** is correct. The PSK is shared between the wireless client and the AP that it is trying to associate with. The key string is not actually passed over the air.
3. **D** is correct. WPA3 Enterprise is considered the most secure form of wireless authentication.
4. **A** is correct. Web passthrough mode is a slight variation of regular web authentication in which the user is not prompted to enter credentials.

# Review Questions

1. Which wireless authentication method can you use with Open Authentication to present an acceptable use policy acknowledgment?
   - ❍ **A.** PSK
   - ❍ **B.** WebAuth
   - ❍ **C.** EAP
   - ❍ **D.** RADIUS

2. Which of the following wireless authentication methods does WPA3 Personal support?
   - ❍ **A.** Open Authentication
   - ❍ **B.** PSK
   - ❍ **C.** EAP
   - ❍ **D.** WebAuth

3. Which wireless authentication method is required if you want to integrate wireless client authentication with 802.1X?
   - ❍ **A.** Open Authentication
   - ❍ **B.** PSK
   - ❍ **C.** EAP
   - ❍ **D.** WebAuth

# Answers to Review Questions

1. **B** is correct. Open Authentication can be used with WebAuth, where the wireless user is prompted on a web page to acknowledge an acceptable use policy before continuing.
2. **B** is correct. Any of the WPA Personal modes supports authentication with PSK.
3. **C** is correct. Wireless client authentication with 802.1X uses EAP.

# Further Reading

- ▶ *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers network security design.

*This page intentionally left blank*

CHAPTER 10

# Network Security Design

**This chapter covers the following official ENCOR 350-401 exam objectives:**

- 5.5 Describe the components of network security design
  - 5.5.a Threat defense
  - 5.5.b Endpoint security
  - 5.5.c Next-generation firewall
  - 5.5.d TrustSec, MACsec

This chapter covers the first four subsections of ENCOR 350-401 Exam Objectives 5.5, and Chapter 11, "Network Access Control," covers the fifth. This chapter looks at network security design for threat defense and examines how to use Cisco's SAFE security framework to help design security solutions for a network. This chapter also looks at the various components of a network security design for an enterprise network that helps to protect, detect, and remediate security threats. Endpoints in a campus environment are vulnerable to security threats such as malware and ransomware. It is therefore important to have a solid network security design that addresses security threats to the endpoints. Also, you must have systems in place that enforce network access control by validating the identities of end users to determine who they are and what they are allowed to access before being granted access to the network. Cisco TrustSec technology, which is covered in this chapter, is one such system. This chapter concludes by looking at how to use Media Access Control Security (MACsec) to provide Layer 2 encryption between switches and between endpoints and switches.

**This chapter covers the following technology topics:**

- Threat Defense
  - Network Security Components
- TrustSec, MACsec

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. In which Cisco SAFE security framework PIN are the majority of a company's information assets found?
2. What is the name of Cisco's software-defined segmentation solution that dynamically organizes endpoints in logical entities called security groups?

### Answers

1. Data center
2. TrustSec

# Threat Defense

Cybersecurity threats such as malware, ransomware, and phishing are constantly evolving and becoming more common. There is not a single solution to protect an organization from all of these threats. To address them in a broader framework, Cisco created SAFE, a security framework that helps you design secure solutions for places in the network (PINs). SAFE organizes security into two broad components:

- PINs
- Secure domains

PINs are the various locations commonly found in a network, and they conceptually represent the infrastructure deployed at those locations. Let's take a look at these PINs:

- **Branch:** Typically, the branch is less secure than the corporate site, with its campus and data center. It is usually cost-prohibitive to duplicate the same security controls you have in the campus and data center at the branch. This is especially true when scaling to a larger number of branches. Branch locations are therefore prime targets and more susceptible to breaches. It is critical to include the right balance of security

capabilities while ensuring a cost-effective design. Here are some of the typical threats mitigated at the branch:

- Endpoint (POS [point of sale])
- Wireless infrastructure exploits (rogue APs, man-in-the-middle attacks)
- Unauthorized/malicious client activity
- Exploitation of trust

- **Campus:** The campus has a large user base and a variety of device types. Due to the large number of security zones (subnets and VLANs), secure segmentation is difficult. Due to the traditionally small number of security controls and the need for guest/partner access, the campus is a prime target for attacks. These are some of the typical threats mitigated in the campus:
  - Phishing
  - Web-based exploits
  - Unauthorized network access
  - Malware propagation
  - BYOD, which creates a larger attack surface and increases the risk of data loss
  - Botnet infestation
- **Edge:** The edge is considered the highest-risk PIN because it is the primary point of ingress for public traffic from the Internet and the main egress point for corporate traffic to the Internet. In today's Internet-based economy, it can be considered the most critical resource for a business. Here are some of the typical threats mitigated at the edge:
  - Web server vulnerabilities
  - Distributed denial-of-service (DDoS) attacks
  - Data loss
  - Man-in-the-middle attacks
- **WAN:** The WAN connects all of the network locations to provide a single point of control and access to network resources. Managing security policies to manage communication can become complex. These are some of the typical threats mitigated in the WAN:
  - Malware propagation
  - Unauthorized network access
  - WAN sniffing and man-in-the-middle attacks

- **Data center:** The data center contains the majority of a company's information assets and intellectual assets and therefore is an important target for potential threats. Data centers typically host thousands of servers, and that by itself provides a challenge when it comes to properly creating and managing security rules to control network access. Here are some of the typical threats mitigated in the data center:
  - Data extraction (data loss)
  - Malware propagation
  - Unauthorized network access (application compromise)
  - Botnet infestation (scrumping), data loss, privilege escalation, and reconnaissance
- **Cloud:** The bulk of the risk from using cloud services relates to the loss of control, lack of trust, shared access, and shadow IT. Service-level agreements (SLAs) with cloud services providers can help alleviate some of these risks by allowing businesses to dictate control of security capabilities in the cloud-provided services. Independent certification risk assessment audits can increase trust. These are some of the typical threats mitigated in the cloud:
  - Web server vulnerabilities
  - Loss of access
  - Virus and malware propagation
  - Man-in-the-middle attacks

Figure 10.1 shows the Cisco SAFE security framework.

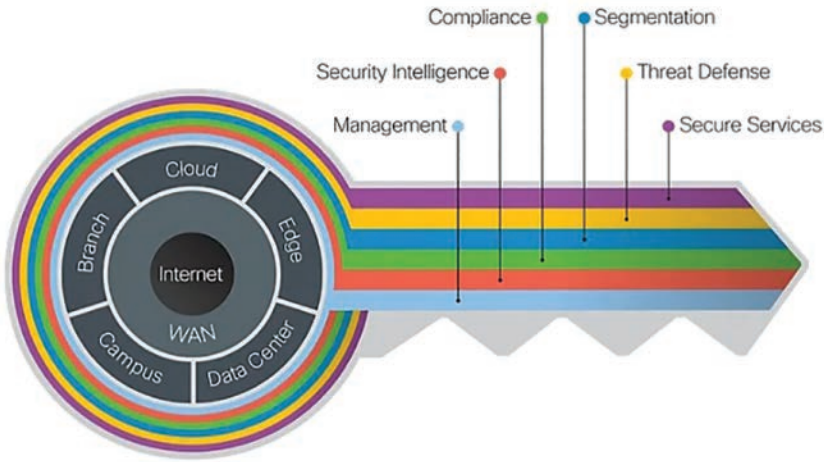



FIGURE 10.1 **Cisco SAFE Security Framework, Showing PINs**

The secure domains represent the operational side of Cisco's SAFE security framework. Operational security is divided by function and the people in the organization who are responsible for each one. Each of the secure domains has class and security capabilities, as well as operational aspects. These security concepts are used to evaluate each PIN.

Let's take a brief look at the secure domains:

- **Secure services:** Secure services enable technologies like access control, VPNs, and encryption. This secure domain also includes protection for insecure services, such as applications, collaboration, and wireless.
- **Threat defense:** Threat defense provides visibility into dangerous cyber threats. By looking at the data from network telemetry, reputation, and contextual information, you can assess the nature of potential risks of a suspicious nature and apply remediation.
- **Segmentation:** Segmentation involves establishing boundaries for data and users. Traditional segmentation methods include use of network addressing, VLANs, and firewall policies for policy enforcement. Advanced segmentation uses identity-aware infrastructure to enforce automated and highly scalable policies.
- **Compliance:** Compliance deals with internal and external policies. It speaks to how multiple controls can be satisfied with a single solution. Examples of external compliance include HIPAA, PCI DSS, and the Sarbanes-Oxley Act (SOX).
- **Security intelligence:** The security intelligence secure domain provides global detection and aggregation of emerging malware. It allows you to enforce policy dynamically, as reputations are augmented by the context of the threats. This in turn gives you accurate and timely security protection.
- **Management:** The management of devices and systems using centralized services is important for consistent policy deployment, workflow change management, and systems patching. The management secure domain coordinates policies, objects, and alerts.

The SAFE framework is designed to be modular, and you can remove a PIN that does not exist in your network when crafting a network security design.

Before closing this section, let's briefly look at the attack continuum and various threat-centric cybersecurity solutions that you can use before, during, and after an attack:

- **Before an attack:** Solutions you can use before an attack include next-generation firewalls, network access control, and identity services. These solutions give you the tools needed to discover threats, enforce policies, and harden existing policies.
- **During an attack:** Solutions you can use during an attack include next-generation intrusion prevention systems (IPSs) and email and web security solutions. These solutions enable you to detect, block, and defend against attacks that have penetrated your system and are in progress.
- **After an attack:** After an attack, you can use solutions such as Cisco Advanced Malware Protection (AMP) and network behavior analysis to determine the scope of the attack and prevent damage.

These solutions, among other security components, are discussed in the next section.

**ExamAlert**

For the ENCOR exam, you need to completely understand the Cisco SAFE security framework and how it aligns with various network security components.

# Network Security Components

This section highlights the various network security components that you can deploy in a network to protect it against attacks. It focuses on the Cisco SAFE security framework, especially the campus PIN. These components are covered in detail in the following subsections:

- Cisco Advanced Malware Protection (AMP)
- Cisco AnyConnect
- Cisco Umbrella
- Cisco Secure Network Analytics
- Content security
  - Cisco Secure Web Appliance
  - Cisco Email Security

- Next-generation IPSs (NGIPS)
- Next-generation firewalls (NGFWs)

## Cisco Advanced Malware Protection (AMP)

The goal of malware in a network is to penetrate a system and avoid detection. Once loaded onto a system, it seeks to self-replicate and insert itself into other programs and files to infect them as well. These threats can take the form of software viruses and other malware, such as ransomware, worms, Trojans, spyware, and adware. Cisco AMP is designed to prevent and detect breaches by advanced malware and aid in removing threats from an infected system. AMP protection is comprehensive in that it protects an organization across the attack continuum described earlier in the chapter (before, during, and after an attack).

One of the robust features of AMP is that it provides comprehensive global threat intelligence. AMP uses Cisco Talos Security Intelligence and Research Group, as well as Threat Grid intelligence feeds. These resources represent an extensive collection of real-time threat intelligence with broad visibility, a large footprint, and the ability to put intelligence into action across multiple security platforms.

AMP can be deployed at various control points throughout the extended network. In addition, a specific AMP solution can be deployed to meet particular security needs. Let's take a look at the AMP components:

- **Cisco AMP for Endpoints:** AMP for Endpoints protects endpoints running Windows, macOS, Linux, and Android. It does this by using a lightweight connector that has no impact on end users. AMP for Endpoints can also be launched from Cisco AnyConnect 4.1. (AnyConnect is covered later in this section.)
- **Cisco AMP for Networks:** Cisco AMP for Networks is a network-based solution integrated into Cisco Firepower NGIPS security appliances.
- **Cisco AMP on ASA firewalls and ASA with Firepower services:** It is possible to deploy AMP capabilities integrated into ASA firewalls or Cisco NGFWs.
- **Cisco AMP Private Cloud Appliance:** It is possible to deploy AMP as an on-premises solution for organizations with privacy requirements that prevent them from using a public cloud.

- **Cisco AMP on ESA or WSA:** AMP capabilities can be used for Cisco Email Security Appliance (ESA) or Web Security Appliance (WSA) to provide retrospective and malware analysis.
- **Cisco AMP for Meraki MX:** It is possible to deploy AMP as part of Meraki MX Security Appliance for cloud-based security management with advanced threat capabilities.
- **Cisco Threat Grid:** Threat Grid is integrated with AMP for enhanced malware analysis. It can also be deployed as a standalone advanced malware analysis and threat intelligence solution, either on an appliance on premises or in the cloud.

Another critical component of AMP is AMP Private Cloud, which does real-time decision-making and is constantly evolving based on the data it receives. It has a database of files and their reputations. This is also known as the file disposition, and it can change in AMP Private Cloud based on Talos or Threat Grid data.

## Cisco AnyConnect

Cisco AnyConnect Secure Mobility Client is modular endpoint software that provides VPN access through Secure Sockets Layer (SSL) and IPsec IKEv2. It offers enhanced security through various modules that offer services such as compliance through the VPN established with ASA. These modules include VPN posture (host scanning), web security, and off-network roaming protection with Cisco Umbrella (covered next).

The modules on the AnyConnect client can scan for software such as antivirus and firewall software to ensure that it is installed and updated on the endpoints. If an endpoint is not in compliance, network access can be restricted until the endpoint gets into compliance. The AnyConnect VPN clients are supported on a broad range of endpoints, including macOS, Windows, Linux, iOS, Android, Windows Phone/Mobile, BlackBerry, and Chrome OS.

## Cisco Umbrella

Cisco Umbrella (formerly known as OpenDNS) enforces security at the DNS and IP layers. It blocks requests to malware, ransomware, phishing, and botnets before a connection is established—basically before the danger reaches an endpoint. Umbrella Secure Web Gateway logs and inspects all web traffic, providing greater control and protection. Also, its cloud-delivered firewall logs and blocks traffic, using IP, port, and protocol rules for consistent enforcement.

Cisco has guaranteed 100% uptime since 2006 with the Umbrella cloud network. The Umbrella global network consists of 30 data centers around the globe. Thanks to anycast routing, these various data centers are available using the same IP address. Requests are transparently sent to the nearest and fastest data center, and failover occurs automatically.

The setup to use Umbrella in a corporate network is straightforward. You simply need to change the DHCP configuration of all of your internet gateways to push out the new DNS configurations. By doing so, you effectively forward all DNS traffic to Umbrella's global network.

## Cisco Secure Network Analytics

Cisco Secure Network Analytics (formerly Stealthwatch) is an agentless solution that uses machine learning and behavioral modeling to detect emerging threats and respond to them quickly. It uses the telemetry data from a network to determine who is on the network and what they are doing. Deployment can be done on premises, in the cloud as a hardware or a virtual appliance, or using Cisco's software-as-a-service (SaaS) offering.

Using a combination of behavioral modeling, machine learning, and global threat intelligence, Cisco Secure Network Analytics can detect threats such as command-and-control (C and C) attacks, ransomware, DDoS attacks, illicit cryptomining, unknown malware, and insider threats. It also has the ability to monitor threats—even on network traffic that is encrypted.

Cisco Secure Network Analytics can be integrated with Cisco Identity Services Engine (ISE) to enforce policies and contain threats. In addition, it is possible to extend visibility to the public cloud. With Secure Cloud Analytics (formerly Stealthwatch Cloud), you can have visibility into the threat detection across all major public cloud platforms.

## Content Security

Content security systems provide granular control and security for specific network applications. In particular, we will take a look at Cisco Secure Web Appliance and Cisco Secure Email here:

- **Cisco Secure Web Appliance:** This product protects a network by automatically blocking risky sites and by testing unknown sites before allowing users to click on them. It uses TLS 1.3 to protect users, and it can be used with Cisco Umbrella to provide comprehensive on-premises

and cloud-based defense against web-based attacks. These are some of the high-level features of Cisco Secure Web Appliance:

- Using Talos threat research, it can do in-depth URL filtering and reputation analysis, Layer 4 traffic monitoring, and malware defense.
- Granular control is achieved by allowing deeper control than just allowing or blocking a website. It can allow or forbid specific functionality with sites but can make exceptions for certain users and set times. It can also do bandwidth restrictions for web usage.
- Flexible deployment options are either hardware or a virtual appliance. It can also be deployed in the public cloud with Amazon Web Services (AWS).
- It can integrate with SecureX to provide enhanced visibility and automation across the Cisco Secure product suite.

- **Cisco Email Security:** This product provides advanced threat protection capabilities to detect, block, and remediate threats and prevent data loss. It can help combat business email compromise (BEC), ransomware, advanced malware, phishing, spam, and data loss through a multilayered approach. Let's look at some of the high-level features of Cisco Email Security:

  - By using Talos threat research, it can detect and block more threats.
  - It can combat ransomware hidden in attachments that evade initial detection with Cisco Secure Endpoint and Cisco Secure Malware Analytics.
  - It can drop emails with risky links automatically or block access to newly infected sites based on real-time URL analysis to protect against phishing and BEC.
  - It prevents brand abuse and sophisticated identity-based email attacks with Cisco Domain Protection (CDP) and Cisco Advanced Phishing Protection (CAPP) services.
  - It protects sensitive content in outgoing emails with data loss prevention (DLP) and easy-to-use email encryption, all in one solution.
  - Cisco Secure Awareness Training provides user behavior training to help users work smarter and safer.
  - It provides flexible deployment options with cloud, virtual, on-premises, and hybrid deployment options.
  - It can be integrated with SecureX to provide enhanced visibility and automation across the Cisco Secure product suite.

## Next-Generation IPSs (NGIPSs)

Cisco Firepower NGIPS security appliances help you mitigate threats by providing deep visibility, preeminent security intelligence, and advanced threat protection for networks. The Firepower NGIPS provides this protection with the ability to operate in-line via the Fail-to-Wire/Bypass network module.

The Firepower NGIPS continuously discovers information about a network, including data about operating systems, mobile devices, files, applications, and users. It uses that information to build network maps and host profiles that give you the contextual information you need to make better decisions about intrusion events. That information is also used as input to better enable the automation of key threat protection features.

The following are some of the key features of a Firepower NGIPS deployment:

- **Talos Security Intelligence and Research Group:** This Cisco group collects and correlates threats in real time. Its efforts result in vulnerability-focused IPS rules and embedded IP, URL, and DNS-based security intelligence for the Firepower NGIPS.
- **Advanced threat protection and rapid remediation:** The Firepower NGIPS can rapidly detect, block, contain, and remediate advanced threats through tight integration with AMP and sandboxing solutions. It can also patch vulnerabilities instantaneously, before new software or signatures become available.
- **Granular application visibility and control:** With the Firepower NGIPS, you can reduce threats to your network through precise control over more than 4000 commercial applications, with support for custom applications.
- **Cisco Firepower Management Center:** This tool provides a single point of event collection and policy management for all deployments of the Cisco Firepower NGIPS, Cisco Firepower Threat Defense for ISR, and the Cisco Firepower NGFW. It gives you a comprehensive enterprise-wide view of security posture, provides consistent security at all points in your network, and reduces management complexity.
- **Integration with other Cisco network security products:** The Firepower NGIPS provides threat effectiveness without complexity. Firepower NGIPS detections can drive automated remediation actions (such as quarantines and blocks) to take place in Cisco's ISE for rapid threat containment.

- **Deployment:** Deployment of the Firepower NGIPS can be done using hardware appliances or via virtualized appliances for the VMware vSphere environment (NGIPSv for VMware).

ExamAlert

For the ENCOR exam, you need to understand the capabilities of an NGFW and the management options for Cisco Firepower.

## Next-Generation Firewalls (NGFWs)

A Cisco NGFW provides capabilities beyond those that are found on traditional stateful firewalls. Apart from doing the traditional stateful inspection of incoming and outgoing traffic, an NGFW includes additional security features, such as application awareness and control, integrated intrusion prevention, and cloud-delivered threat intelligence. An NGFW can also block threats like advanced malware and application-layer attacks. According to Gartner (see https://www.gartner.com/en/information-technology/glossary/next-generation-firewalls-ngfws), an NGFW must meet these capabilities:

- Standard firewall capabilities, such as stateful inspection
- Integrated intrusion prevention
- Application awareness and control to see and block risky apps
- Threat intelligence sources
- Upgraded paths to include future information feeds
- Techniques to address evolving security threats

Cisco's Firepower NGFW integrates ASA firewall software with Firepower NGIPS services—and it exceeds Gartner's NGFW definition. Firepower is a fully integrated, threat-focused NGFW with a host of management features. Apart from the traditional management tools for managing Cisco ASA firewalls, the Cisco Firepower NGFW also provides the following management options:

- **Cisco Defense Orchestrator:** Defense Orchestrator is a cloud-based application that helps you consistently manage policies across Cisco firewall and public cloud infrastructure. Some of the feature highlights of Cisco Defense Orchestrator are:

- Management of security policy consistently across ASA, Cisco Firepower, Meraki MX, and AWS from a centralized console
- Simplified upgrades for ASA and FTD software images that require only a few clicks
- Ability to track every change and continuously document the changes in a viewable change log
- Ability to integrate with other Cisco Security solutions

- **Cisco Security Analytics and Logging:** Security Analytics and Logging enables you to streamline decision-making by aggregating logs from various Cisco devices and providing an intuitive view of network activity.
- **Secure Firewall Management Center:** Secure Firewall Management Center (formerly Firepower Management Center) is the administrative center for Cisco security products running different platforms. It is the centralized event and policy manager for the following:
  - Cisco Secure Firewall with Firewall Threat Defense (FTD)
  - Cisco ASA with Firepower Services
  - NGIPSs
  - Cisco Firepower Threat Defense for ISR
  - Cisco Malware Defense (AMP)

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. What is the name of Cisco security framework?
   - ❍ **A.** Cisco Validated Design
   - ❍ **B.** SAFE
   - ❍ **C.** Secure Domain
   - ❍ **D.** Threat Grid

2. Which PIN is considered the highest risk PIN in Cisco's SAFE security framework?
   - ❍ **A.** WAN
   - ❍ **B.** Data center
   - ❍ **C.** Edge
   - ❍ **D.** Branch

3. Which Cisco security solution is agentless and uses machine learning and behavioral modeling to detect emerging threats?
   - ❍ **A.** Cisco AMP
   - ❍ **B.** Cisco Umbrella
   - ❍ **C.** Cisco Secure Network Analytics
   - ❍ **D.** Cisco AnyConnect

# Answers

1. **B** is correct. Cisco SAFE is a security framework that Cisco created to help design secure solutions for places in the network (PINs).
2. **C** is correct. The edge is considered the highest-risk PIN because it is the primary point of ingress for public traffic from the Internet and the main egress point for corporate traffic to the Internet.
3. **C** is correct. Cisco Secure Network Analytics (formerly Stealthwatch) is an agentless solution that uses machine learning and behavioral modeling to detect emerging threats and respond quickly.

# TrustSec, MACsec

This section covers TrustSec and MACsec technology.

**ExamAlert**

For the ENCOR exam, make sure you understand TrustSec, including SGT, SGACLs, and its three phases.

## TrustSec

TrustSec is a Cisco software-defined segmentation solution that dynamically organizes endpoints in logical entities, called *security groups*. These security groups are assigned based on business decisions using a richer context than an IP address since they are easier for people to understand and manage.

The TrustSec technology is embedded in a lot of Cisco and third-party products. Using microsegmentation, TrustSec isolates attacks and quickly restricts the lateral movement of threats. It also enables a scalable bring-your-own-device (BYOD) environment and reduces the scope of compliance for industry and government regulations.

A Cisco TrustSec policy group called a Security Group Tag (SGT) is assigned to an endpoint at the network access entry point. This is typically based on that endpoint's user, device, and location attributes. The SGT denotes the endpoint's access entitlements, and all traffic from the endpoint will carry the SGT information as a tag in the frame. Switches, routers, and firewalls use the SGT to make forwarding decisions. Because SGT assignments can denote business roles and functions, TrustSec controls can be defined in terms of business needs and not underlying networking constructs, as is the case with VLANs.

Let's now look at some of the main highlights and benefits of a Cisco TrustSec deployment in an environment:

- Policy can be implemented using security groups ACL (SGACLs) and can be provisioned dynamically or statically on routers, switches, and wireless LAN infrastructure.

  TrustSec integrates with ISE, which acts as the controller for software-defined segmentation groups and policies. This provides a layer of policy abstraction and centralized administration.

- The integration with ISE allows for simplification of cross-domain security policy. You can share TrustSec group information with other group-based policy schemes in Cisco's Application-Centric Infrastructure (ACI).

TrustSec is defined as having three phases:

1. **Classification:** When users or devices connect to the network, they are assigned to security groups. This is the classification phase, and it can be based on the results of authentication or can occur by associating the SGT with an IP, VLAN, or port profile. Classification can happen in one of two ways:
   - **Dynamic assignment:** The SGT assignment is done dynamically and can be downloaded as an authentication option from ISE when authenticating with 802.1x, MAB, or WebAuth (covered in Chapter 11).
   - **Static Assignment:** The SGT assignment is done statically when dynamic assignment is not possible. The SGT tags can be statically mapped on SGT-capable network devices.
2. **Propagation:** Once the traffic is classified, the SGT is propagated from where the classification took place to where enforcement action is invoked. There are two methods of SGT propagation:
   - **Inline tagging:** SGT is embedded into the Ethernet frame. Embedding the SGT into the Ethernet frame requires specific hardware support.
   - **Use of SGT Exchange Protocol (SXP):** SXP is used to share the SGT-to-IP address mapping. This allows for SGT propagation to continue to the next device in the path. Network devices do not have to have specific hardware support to use SXP.
3. **Enforcement:** During the enforcement phase, traffic is controlled based on the tag information. The enforcement point can be a Cisco router, switch, or firewall. The enforcement device takes the source SGT and looks it up against the destination SGT and determines if the traffic should be allowed or denied.

One of the critical pieces of TrustSec implementation is ISE (which is covered in Chapter 11). ISE is commonly used as the centralized repository for SGT, security groups, and SGACLs. For example, tags, such as Employees_SGT and Dev Servers_SGT, can represent the user group Employees and the server group Dev Servers. These tags are then used as sources and/or destinations in an access policy.

# MACsec

The demand for increased bandwidth continues, driven mainly by cloud services, mobile devices, and a tremendous increase in video traffic. The shift to cloud and mobile service accelerates the need for faster WAN transport speeds to handle the traffic related to locating applications and data off premises. With the increased demand for high-speed links comes the need to protect these links, and this is where the 802.1AE Media Access Security (MACsec) technology comes into play.

> **ExamAlert**
>
> For the ENCOR exam, make sure you have a complete understanding of MACsec and how it differs from IPsec.

MACsec is a MAC layer or link-layer encryption technology that enables point-to-point encryption between two MACsec peers. Unlike IPsec, which is typically used on a central application-specific integrated circuit (ASIC) optimized for encryption, MACsec is enabled on a per-port basis, leveraging onboard ASICs. MACsec is based on standard Ethernet frame format, but a 16-byte MACsec security tag (SecTAG) is included. In addition, a 16-byte Integrity Check Value (ICV) is included at the end of the frame.

The 16-byte SecTAG field is as follows:

- **MACsec EtherType (octets 1 and 2):** The first two octets and the value are set to 0x88e5 and designate that the next frame is a MACsec frame.
- **TCI/AN (octet 3):** The third octet is the Tag Control Information (TCI)/Association Number field. The TCI designates the MACsec version number.
- **SL (octet 4):** The fourth octet is Short Length, which is set to the length of the encrypted data.
- **PN (octets 5–8):** Octets 5 through 8, the Packet Number octets, are used for replay protection and the construction of the initialization vector along with the secure channel identifier (SCI).
- **SCI (octets 9–16):** Octets 9 through 16 are the Secure Channel Identifier octets. Each connectivity association (CA) is a virtual port. Each virtual port is designated a secure channel identifier that is the combination of the MAC address of the physical interface and a 16-bit port ID.

A MACsec frame can be encrypted and authenticated to provide both privacy and integrity. It uses Galois/Counter Mode Advanced Encryption Standard (AES-GCM) for authenticated encryption. It uses Galois Message Authentication Code (GMAC) if only authentication, and not encryption, is required.

Depending on what links are being encrypted, there are two implementations of MACsec:

- **Downlink MACsec:** This type of MACsec encrypts the link between an endpoint and a switch. This encryption is handled by the MAC security key agreement (MKA) keying protocol. This implementation requires a MACsec-capable Catalyst switch and a MACsec-capable supplicant on the endpoint (for example, Cisco AnyConnect).
- **Uplink MACsec:** This type of MACsec encrypts the link between switches with the 802.1AE standard. The switch-to-switch encryption uses the Cisco-proprietary Security Association Protocol (SAP). The encryption uses AES-GCM-128 encryption, which is used by both downlink and uplink MACsec.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which TrustSec phase controls traffic based on the tag information?
   - ❍ **A.** Classification
   - ❍ **B.** Propagation
   - ❍ **C.** Enforcement
   - ❍ **D.** Assignment

2. Which TrustSec phase assigns users and devices to security groups?
   - ❍ **A.** Classification
   - ❍ **B.** Propagation
   - ❍ **C.** Enforcement
   - ❍ **D.** Assignment

3. True or false: MACsec is implemented on a per-port basis.
   - ❍ **A.** True
   - ❍ **B.** False

# Answers

1. **C** is correct. Enforcement controls traffic based on the tag information. The enforcement point can be a Cisco router, switch, or firewall.
2. **A** is correct. Classification happens when the users or devices connect to the network. They are assigned to security groups. Classification is based on authentication results or associating the SGT with an IP address, VLAN, or port profile.
3. **A** is correct. MACsec is enabled on a per-port basis, leveraging onboard ASICs.

# Review Questions

1. Which of the SAFE secure domains speaks to global detection and aggregation of emerging malware?
   - ❍ **A.** Threat defense
   - ❍ **B.** Compliance
   - ❍ **C.** Segmentation
   - ❍ **D.** Security intelligence

2. Which of the following is a Cisco security solution that blocks requests to malware, ransomware, phishing, and botnets before they reach an endpoint at the IP and DNS layer?
   - ❍ **A.** Cisco AMP
   - ❍ **B.** Cisco Umbrella
   - ❍ **C.** Cisco Secure Network Analytics
   - ❍ **D.** Cisco AnyConnect

3. In which TrustSec classification method is the assignment of SGT done seamlessly and downloaded from ISE when authenticating with 802.1x, MAB, or WebAuth?
   - ❍ **A.** Dynamic assignment
   - ❍ **B.** Static assignment
   - ❍ **C.** Inline tagging
   - ❍ **D.** SXP

4. True or false: Downlink MACsec is implemented by encrypting the links between switches.
   - ❍ **A.** True
   - ❍ **B.** False

# Answers to Review Questions

1. **D** is correct. The security intelligence secure domain provides for global detection and aggregation of emerging malware.
2. **B** is correct. Cisco Umbrella blocks requests to malware, ransomware, phishing, and botnets before a connection is established—basically before it reaches an endpoint.

3. **A** is correct. With dynamic assignment, the assignment of SGT is done dynamically and can be downloaded as an authentication option from ISE when authenticating with 802.1x, MAB, or WebAuth.
4. **B** is correct. Downlink MACsec is the encrypting of the link between an endpoint and the switch.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers network access control.

*This page intentionally left blank*

CHAPTER 11

# Network Access Control

**This chapter covers the following official ENCOR 350-401 exam objectives:**

- 5.5 Describe the components of network security design
- 5.5.e Network Access Control with 802.1X, MAB, and WebAuth

Chapter 10, "Network Security Design," cover the first four subsections of ENCOR 350-401 Exam Objective 5.5, and this chapter covers the fifth and final subsection of that exam objective. This chapter starts by looking at Cisco Identity Services Engine (ISE), including the features, benefits, and integrations that ISE supports. Then this chapter looks at how to use 802.1X, MAB, and WebAuth to implement network access control in a network environment.

**This chapter covers the following technology topics:**

- Cisco Identity Services Engine (ISE)
  - Network Access Control (NAC)

## Cram Saver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. Which authentication mechanism is typically used as a fallback for 802.1X when authenticating devices?
2. Which WebAuth type supports advanced services such as client provisioning, posture assessments, acceptable use policies, self-registration, and device registration?

### Answers

1. MAC Authentication Bypass (MAB)
2. Central Web Authentication

# Cisco Identity Services Engine (ISE)

Cisco Identity Services Engine (ISE) is a security policy management platform that provides secure access to users and devices across wired, wireless, and VPN connections to a corporate network. It is part of Cisco's network access control (NAC) solution and allows you to gain visibility into what is happening on your network. For example, it allows you to see who is connected to your network and what applications are installed and running.

These are some of the high-level features and benefits of Cisco ISE:

- **Centralized management:** ISE has a single web GUI console to centrally configure and manage profiler, posture, guest, authentication, and authorization services.
- **Contextual identity and business policy:** ISE provides a rule-based, attribute-driven policy for flexible business-relevant access control policies. Attributes include user and endpoint identity and the authentication protocol. ISE can integrate with identity repositories such as LDAP, RADIUS, and Active Directory.
- **Access control:** ISE provides a range of access control options, including downloadable access control lists (dACLS), VLANs assignment, URL redirection, and security group ACLs (SGACLs) with Cisco TrustSec.

- **TrustSec:** ISE facilitates software-defined segmentation through the use of a Security Group Tag (SGT) by serving as the policy controller for Cisco TrustSec.
- **Guest life cycle management:** ISE provides a streamlined experience for implementing and customizing guest network access. For guest network access, ISE tracks access across the network for surety, compliance, and auditing. For security control, it facilitates the management of time limits, account expirations, and SMS verifications.
- **Streamlined device onboarding:** ISE automates supplicant provisioning and certificate provisioning for standard PC and mobile computing platforms.
- **Built-in AAA support:** ISE supports a standard RADIUS protocol for authentication, authorization, and accounting (AAA). It supports a wide range of authentication protocols, including PAP, CHAP, EAP-MD5, PEAP, EAP-Flexible Authentication via Secure Tunneling (FAST), EAP-Transport Layer Security (TLS), and EAP-Tunneled Transport Layer Security (TTLS).
- **Device profiling:** ISE is populated with predefined device templates for many different types of endpoints, including IP phones, IP cameras, smartphones, and tablets.
- **Device-profile feed service:** ISE can deliver automatic updates of Cisco's validated device profiles for various IP-based devices from different vendors.
- **Extensive Active Directory support:** ISE provides comprehensive authentication and authorization against multiforest Microsoft Active Directory domains. It supports Microsoft Active Directory 2003, 2008, 2008 R2, 2012, 2012 R2, 2016, and 2019.
- **Certification:** ISE meets the requirements for Federal Information Processing Standard (FIPS) 140-2, Common Criteria, and Unified Capabilities Approved Product List.
- **IPv6 support:** ISE supports IPv6 for RADIUS- and TACACS+-based network devices.
- **Platform support and compatibility:** ISE is available as both a physical appliance and a virtual appliance. Both deployments can create an ISE cluster to scale, provide redundancy, and provide failover requirements in critical network environments. The virtual appliance is supported on VMware ESXi 5.x and 6.x, KVM on Red Hat 7.x, and Microsoft Hyper-V on Microsoft Windows Server 2012 R2 and later.

# Network Access Control (NAC)

A network access control (NAC) solution typically supports network visibility and access management through some form of policy enforcement on network devices and users. NAC solutions generally have the following capabilities:

- **Policy life cycle management:** A NAC solution enforces policies for all operating scenarios without requiring separate products or additional modules.
- **Profiling and visibility:** A NAC solution recognizes and profiles users and their devices before malicious code can cause harm in a network.
- **Guest networking access:** A NAC solution makes it possible to manage guests through a customizable, self-service portal that includes guest registration and guest authentication.
- **Security posture check:** A NAC solution evaluates security policy compliance by user type, device type, and operating system.
- **Incidence response:** A NAC solution mitigates network threats by enforcing security policies that block, isolate, and repair noncompliant machines without requiring administrator attention.
- **Bidirectional integration:** A NAC solution integrates with other security and network solutions through an open/RESTful API.

Cisco ISE is part of Cisco's NAC solution that enables guests/contractors to make sure that non-employees have access privileges to the network that are separate from those of employees. Cisco's ISE NAC solution also provides secure bring-your-own-device (BYOD) access to ensure compliance for employee-owned devices before they access the network.

Next, we look at network access control with 802.1X, MAB, and WebAuth. Cisco TrustSec and MACsec are covered in Chapter 10.

**ExamAlert**

For the ENCOR exam, make sure you understand these components of NAC: 802.1X, MAB, and WebAuth.

# 802.1X

802.1X (pronounced "dot 1 X") is a client/server-based network access control mechanism that restricts unauthorized clients (supplicants) from accessing the

network. An authentication server (RADIUS server) validates each supplicant connected to an authenticator (a network access device such as a switch or WLAN controller) before making any service offered by the network access device available.

Let us look more closely at the different device roles in an 802.1X deployment:

- **Client (supplicant):** The client requests access to the network and responds to requests from the network access device. The client runs an 802.1X-compliant client software. Some supplicants include the native Windows and native macOS supplicants and Cisco AnyConnect.
- **Authenticator:** The authenticator controls physical access to the network, based on the status of the authentication of the client. The network access devices serve as authenticators and are intermediaries between the authentication server and the client. A network access device requests identity information from the client, verifies that information with the authentication server, and then relays the response to the client. The network access device works by encapsulating and decapsulating the EAP frame and interacting with the RADIUS server.
- **Authentication server:** The authentication server (for example, the RADIUS server) performs the authentication of the client. It validates the client's identity and notifies the network access device that the client is authorized to access the network.

Let's now take a look at the port-based authentication process using 802.1X on a Cisco Catalyst switch:

1. To configure IEEE 802.1X port-based authentication, you must enable AAA and specify the authentication method list. The AAA process starts with authentication. If 802.1X port-based authentication is enabled and the client supports 802.1X-compliant client software, you can move through the remaining stages.
2. If the client identity is valid and the 802.1X authentication succeeds, the switch grants the client access to the network.
3. If 802.1X authentication times out while waiting for an EAPOL message exchange and MAC authentication bypass is enabled, the switch can use the client MAC address for authorization. If the client MAC address is valid and the authorization succeeds, the switch grants the client access to the network.

4. If the switch gets an invalid identity from an 802.1X-capable client and a restricted VLAN is specified, the switch can assign the client to a restricted VLAN that provides limited services.
5. If the RADIUS authentication server is unavailable (down) and inaccessible authentication bypass is enabled, the switch grants the client access to the network by putting the port in the critical-authentication state in the RADIUS-configured or user-specified access VLAN.

Let us now look at the steps for port-based authentication initiation and message exchange:

1. The client or Catalyst switch initiates authentication. When the switch initiates authentication, the switch sends an EAP-request/identity frame to the client to request its identity. Upon receipt of the frame, the client responds with an EAP-response/identity frame.
2. If during bootup the client does not receive an EAP-request/identity frame from the switch, the client can initiate authentication by sending an EAPOL-start frame, which prompts the switch to request the client's identity.
3. When the client provides its identity, the switch begins its role as the intermediary, passing EAP frames between the client and the authentication server until authentication succeeds or fails. If the authentication succeeds, the switch port becomes authorized.
4. If the authentication fails, authentication can be retried, or the port might be assigned to a VLAN that provides limited services or network access is denied.
5. If 802.1X authentication times out while waiting for an EAPOL message exchange and MAC Authentication Bypass is enabled, the switch can authorize the client when the switch detects an Ethernet packet from the client.

## MAC Authentication Bypass (MAB)

MAC Authentication Bypass (MAB) is a MAC address–based authentication mechanism that enables port-based access control using the MAC address of the endpoint. The use of MAB is applicable to the following network environments:

- In network environment where a supplicant code is not available for a given client platform

- In a network environment where the endpoints are not under administrative control or where 802.1X requests are not supported on the network

> **ExamAlert**
>
> For the ENCOR exam, make sure you know when MAB is used instead of 802.1X.

MAB is typically used as a fallback mechanism to 802.1X and is ideal for situations where you cannot use 802.1X because a supplicant is not supported on the end device but you still want to secure the switch port for a network printer, IP camera, or some other network device. Once MAB is enabled on a switch port, the switch drops all frames except for the initial few to learn the MAC address. Once the MAC address is learned, the RADIUS server is queried to check whether it permits the MAC address.

Because a MAC address can be easily spoofed (with an endpoint configured to use a different address from the burned-in one), it is not a secure authentication option. Therefore, the MAB authenticated endpoints should be given restricted access and should only be allowed to communicate with network services that the endpoint actually needs to communicate with. If the authenticator is a Catalyst switch, a few authorization options can be applied as part of the authorization results from the RADIUS server:

- SGT tags
- dACLs
- Dynamic VLAN assignment (dVLAN)

Finally, you can configure MAB for two types of scenarios:

- **Standalone:** With a standalone configuration, you use MAB alone for authentication.
- **Fallback:** With a fallback configuration, you fall back to using MAB after 802.1X is attempted and the attempt is unsuccessful.

## WebAuth

> **ExamAlert**
>
> For the ENCOR exam, make sure you know the use case for WebAuth and when it is used rather than 802.1X or MAB.

WebAuth allows you to control network access and enforce policy based on the authenticated identity of users. It helps you mitigate unauthorized access and provides network access for end hosts that do not support 802.1X. An ideal use of WebAuth is for contractors and consultants or visitors who need access to the Internet. 802.1X does a great job securing an internal network by requiring users to present valid credentials before accessing the network. WebAuth fills the gap when you have users without 802.1X supplicants or where you may not know the MAC address for performing MAB.

Like MAB, WebAuth can be used as a fallback authentication mechanism for 802.1X. If both MAB and WebAuth are set to be used as a fallback for 802.1X and 802.1X times out, a switch will first try to use MAB. If authentication through MAB fails, then the switch will attempt to authenticate through WebAuth. This automatic sequencing of authentication methods allows a network administrator to apply the same configuration to every switch port without knowing in advance the kind of device (for example, guest, printer, IP camera, 802.1X capable or not) that will be connecting to the network.

When the switch falls back to WebAuth, WebAuth authenticates the user by providing a web-based login page on which the user can enter credentials. The credentials are submitted from the switch or other network access device to the RADIUS server. After the user is identified, the user's identity can be used in mapping identities to policies that grant or deny network access.

There are two types of WebAuth:

- **Local Web Authentication:** A switch or another network access device redirects web traffic to a locally hosted web portal running on the network access device where an end user provides his or her credentials. Local Web Authentication has some limitations. It supports ACL assignments and does not support VLAN assignments. In addition, there is no support for the change of authorization (CoA) feature, so access policy cannot be changed based on posture and profiling states.
- **Central Web Authentication with Cisco ISE:** A central device acts as a web portal (ISE in this case). The major difference compared to the Local Web Authentication is that Central Web Authentication is shifted to Layer 2, along with MAC/802.1X authentication. In addition, the RADIUS server returns special attributes that indicate to the switch that a web redirection must occur. This solution has the advantage of eliminating any delay that was necessary for WebAuth to kick in.

- Central Web Authentication supports the CoA feature to apply new policies so that access policies can be changed based on posture and profiling states. It also supports advanced services such as client provisioning, posture assessments, acceptable use policies, self-registration, and device registration.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. True or false: ISE provides support for IPv6 for RADIUS- and TACACS+-based network devices.
   - ❍ **A.** True
   - ❍ **B.** False

2. Which component of the 802.1X infrastructure serves as the intermediatory device and handles identity information from clients?
   - ❍ **A.** Supplicant
   - ❍ **B.** Authenticator
   - ❍ **C.** Authentication server
   - ❍ **D.** TACACS+

3. True or false: MAB can work with 802.1X authentication and can also serve as a fallback for it.
   - ❍ **A.** True
   - ❍ **B.** False

## Answers

1. **A** is correct. ISE supports IPv6 for both RADIUS and TACACS+.
2. **B** is correct. The authenticator serves as the intermediary between the authentication server and the client or supplicant.
3. **A** is correct. MAB is a MAC address–based authentication mechanism that enables port-based access control using the MAC address of the endpoint. It is typically used as a fallback mechanism to 802.1X.

# Review Questions

1. Which component of the 802.1X infrastructure requests access to the network and responds to requests from the network access device?
   - ❍ **A.** Supplicant
   - ❍ **B.** Authenticator
   - ❍ **C.** Authentication server
   - ❍ **D.** TACACS+

2. True or false: WebAuth cannot be used as a fallback authentication mechanism if you are already using MAB as a fallback authentication mechanism for 802.1X.
   - ❍ **A.** True
   - ❍ **B.** False

3. Which authentication mechanism allows for secure network access when you may not know a client's MAC address or when a client does not have a supplicant?
   - ❍ **A.** 802.1X
   - ❍ **B.** MAB
   - ❍ **C.** WebAuth
   - ❍ **D.** RADIUS

# Answers to Review Questions

1. **A** is correct. The supplicant requests access to the network and responds to requests from the network access device.
2. **B** is correct. Like MAB, WebAuth can be used as a fallback authentication mechanism for 802.1X. This is true even if MAB is already being used.
3. **C** is correct. WebAuth fills the gap when you have users without 802.1X supplicants or when you may not know the MAC address to perform MAB.

# Further Reading

- ▶ *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers the anatomy of Python.

*This page intentionally left blank*

CHAPTER 12

# Anatomy of Python

**This chapter covers the following official ENCOR 350-401 exam objective:**

- 6.1 Interpret basic Python components and scripts

Python is one of the most popular, easy-to-learn, and powerful programming languages. This chapter looks at Python components and scripts and how to interpret them. It covers the high-level steps involved in using Python and the steps that are necessary to set up its environment. It also looks at interpreting Python scripts that are used to simplify network automation.

**This chapter covers the following technology topic:**

- Interpreting Python Components and Scripts

# Interpreting Python Components and Scripts

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. Different Cisco platforms support Python in two modes. What are the modes?
2. What is the **type ()** command used for?

### Answers

1. Interactive and noninteractive (script)
2. The **type ()** command returns the type of object.

## Python Overview

Historically, network engineers have manually configured network devices by using the command-line interface (CLI), which is both laborious and prone to human error. The demand for network and service availability across the spectrum has changed, and so has network infrastructure. Network engineers must now mitigate downtime and network outages to meet the demand for higher network and service availability through event-driven actions. Trying to provide high network and service availability by manually performing tasks is not feasible, scalable, or cost-effective. For today's critical network infrastructure that organizations heavily depend on, automation and programming of network devices are critical to maintaining high network and service availability.

Python, which was created in 1991, is an easy-to-use, easy-to-learn, and powerful programming language that allows you to automate networks. Python has efficient high-level data structures, and it takes a simple but powerful approach to object-oriented programming. Python has elegant syntax and an interpreted nature, which make it an ideal language for scripting and rapid application development in many areas on most platforms. Its syntax is similar to that of the English language, so it is easy for a network engineer with little or no programming background to learn. Python is well suited for use in network troubleshooting, general usage, event-based actions, and other DevOps operations.

# Python Releases

Currently, two main releases of Python are used: 2.7 and 3.7. Release 2.7 is more widely used. Cisco switching platforms like the Catalyst 3850 and 9300 support Release 2.7. Because Release 2.7 is installed on some Cisco platforms by default, this release is called *native*, or *onboard*, *Python*. IOS XE Amsterdam 17.3.1 and later releases support only Python 3. Different Cisco platforms support Python in interactive and noninteractive (script) modes. You can invoke Python in the interactive mode in the CLI by running **guestshell run python**. (You will learn more about Guest Shell shortly.) You can run a Python script in noninteractive mode by providing the Python script name as an argument to the **python** command.

Table 12.1 shows the high-level differences between Python 2.x and 3.x.

TABLE 12.1 **Differences Between Python 2.x and 3.x**

| Python 2.x | Python 3.x |
|---|---|
| No longer in active development | Recommended version to use |
| Designed with better library support | Designed to be easier to learn |
| Supported on many Cisco platforms | Supported on newer platforms (for example, Catalyst 9300 running IOS XE Amsterdam 17.3.1 and later releases) |
| Default on Linux and macOS | Can be used with the Python virtual environment (virtualenv) |

# Setting Up Guest Shell

You need to set up the Guest Shell environment before you can use Python interactive mode. Guest Shell is a virtualized Linux-based environment that is designed to run custom Linux applications, including Python, for automated control and management of Cisco devices. This container shell allows you to install scripts or software packages and run them. Commands that are executed through Guest Shell are executed with the same privilege that a user has when logged into the IOS terminal.

Guest Shell is part of the broader IOx application framework. IOx is a Cisco-developed end-to-end application framework that provides application hosting capabilities for different applications types on Cisco platforms; Guest Shell is one such application.

Example 12.1 shows the configuration and verification of the IOx service (which is a prerequisite for enabling and running Guest Shell). You configure IOx in global configuration mode by using the command **iox**. You verify the configuration by using the command **show iox-service**. (The details of the inner workings of IOx and its services are beyond the scope of the ENCOR exam.) You enable Guest Shell using the command **guestshell enable** and then verify it by using the command **show app-hosting list.**

EXAMPLE 12.1 **Configuring IOx and Guest Shell**

```
C3850-01#
C3850-01# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
C3850-01(config)# iox
C3850-01(config)# exit
C3850-01# show iox-service

IOx Infrastructure Summary:
---------------------------
IOx service (CAF)    : Running
IOx service (HA)     : Running
IOx service (IOxman) : Running
Libvirtd             : Running

C3850-01# guestshell enable
Interface will be selected if configured in app-hosting
Please wait for completion
guestshell activated successfully
Current state is: ACTIVATED
guestshell started successfully
Current state is: RUNNING
Guestshell enabled successfully
C3850-01#
C3850-01# show app-hosting list
App id                           State
------------------------------------------------------
guestshell                       RUNNING
```

# Using Python

Example 12.2 shows how to examine the version of Python running on the Cisco Catalyst 3850 platform. You can run Python interactively, or you can run Python scripts through Guest Shell. To verify the Python version in use, you

use the command **guestshell run python**. From the output in this example, you can see that Python Version 2.7.11 is running here.

EXAMPLE 12.2 **Verifying the Python Version**

```
C3850-01#
C3850-01# guestshell run python
Python 2.7.11 (default, May 23 2019, 07:28:05)
[GCC 5.3.0] on linux2
Type "help", "copyright", "credits" or "license" for more information.
>>>
```

**ExamAlert**

For the ENCOR exam, you need to know the high-level tasks that can be performed using Python scripting capabilities.

Python enables you to perform various tasks, such as the following:

- Run a script to verify the configuration on switch bootup
- Back up a configuration
- Perform proactive congestion management by monitoring and responding to buffer utilization characteristics
- Integrate with the Embedded Event Manager (EEM) module
- Perform a job at a particular time interval
- Access the CLI to perform various tasks

You can execute Python code in two ways:

- Using the Python dynamic interpreter
- Writing Python scripts

Example 12.3 shows how to enter and exit the Python shell. In Example 12.2, you saw that executing **guestshell run python** at the switch terminal takes you into the Python shell. You exit the Python shell by using **exit ()** or Ctrl+D. This takes you back to privileged EXEC mode. Because Python is built into the switching platform, no text editor or integrated development environment (IDE) is needed. This helps give Python its lower barrier to entry compared to other programming languages.

EXAMPLE 12.3 **Entering and Exiting the Python Shell**

```
C3850-01#
C3850-01# guestshell run python
Python 2.7.11 (default, May 23 2019, 07:28:05)
[GCC 5.3.0] on linux2
Type "help", "copyright", "credits" or "license" for more information.
>>> exit()
C3850-01#
```

As you saw earlier in this chapter, with the interactive interpreter or Python shell, you can write and test code without writing a full program or script. Also, Python is considered a dynamic language, meaning it does not require a variable (that is, a string) to be defined before it can be used. For example, if you are creating a new variable **C9300**, the only syntax needed is:

```
>>> hostname = 'C9300'
```

Scripts can be written locally or written in a text editor and stored with the **.py** extension. A script can be invoked with the command **guestshell run python /flash/***script_name.py*.

Example 12.4 shows how to run a Python script from flash (that is, non-interactively). The script name in this case is examcram_script.py.

EXAMPLE 12.4 **Running a Python Script**

```
C3850-01# guestshell run python /flash/examcram_script.py
```

Cisco provides a Python module that enables you to run EXEC and configuration commands. You can display the details of the Cisco Python module by entering the **help()** command.

Example 12.5 shows how to view information related to Cisco's Python module. You can use the **help ()** command to execute one of the six functions used to execute CLI commands. This example uses **help()** for the **cli** function.

EXAMPLE 12.5 **Cisco's Python Module Information**

```
C3850-01#
C3850-01# guestshell run python
Python 2.7.11 (default, May 23 2019, 07:28:05)
[GCC 5.3.0] on linux2
Type "help", "copyright", "credits" or "license" for more information.
>>>
>>> from cli import cli,clip,configure,configurep, execute, executep
```

```
>>> help(cli)
Help on function cli in module cli:

cli(command)
    Execute Cisco IOS CLI command(s) and return the result.

    A single command or a delimited batch of commands may be run. The
delimiter is a space and a semicolon, " ;". Configuration commands
must be in fully qualified form.

    output = cli("show version")
    output = cli("show version ; show ip interface brief")
    output = cli("configure terminal ; interface gigabitEthernet 0/0 ;
no shutdown")

    Args:
        command (str): The exec or config CLI command(s) to be run.

    Returns:
        string: CLI output for show commands and an empty string for
            configuration commands.

    Raises:
        errors.cli_syntax_error: if the command is not valid.
        errors.cli_exec_error: if the execution of command is not
successful.
```

Python uses six functions that can execute CLI commands. These functions are available from the Python CLI module. To use these functions, you execute the **import cli** command. Arguments for these functions are strings of CLI commands. To execute a CLI command through the Python interpreter, you enter the CLI command as an argument string of one of these six functions:

- **cli.cli(command):** This function takes an IOS command as an argument, runs the command through the IOS parser, and returns the resulting text.
- **cli.clip(command):** This function works exactly the same as the **cli.cli(command)** function, except that it prints the resulting text to **stdout** rather than return it.
- **cli.execute(command):** This function executes a single EXEC command and returns the output.
- **cli.executep(command):** This function executes a single command and prints the resulting text to **stdout** rather than return it.

- **cli.configure(command):** This function configures the device with the configuration available in commands.
- **cli.configurep(command):** This function works exactly the same as the **cli.configure(command)** function, except that it prints the resulting text to **stdout** rather than return it.

Example 12.6 shows the execution of two CLI commands through the Python interpreter: **cli.configure** and **cli.configurep**. Notice that the command parameters can be multiple lines and in the same format that is displayed with the **show running-config** command.

EXAMPLE 12.6 **Python Module IOS CLI Commands**

```
>>> cli.configure(["interface GigabitEthernet1/0/1", "no shutdown",
"end"])
[ConfigResult(success=True, command='interface GigabitEthernet1/0/7',
line=1, output='', notes=None), ConfigResult(success=True, command='no
shutdown',
line=2, output='', notes=None), ConfigResult(success=True,
command='end',
line=3, output='', notes=None)]
>>>
>>> cli.configurep(["interface GigabitEthernet1/0/1", "no shutdown",
"end"])
Line 1 SUCCESS: interface GigabitEthernet1/0/1
Line 2 SUCCESS: no shutdown
Line 3 SUCCESS: end
>>>
```

There are a number of helper utilities and functions in Python:

- **help ():** This command interprets the built-in documentation about an object.
- **dir ():** This command returns the attributes and methods of an object or module. When used with an argument, it shows the entire list of attributes in the current global scope.
- **type ():** This command returns the type of object.

Python uses a number of data types:

- **String:** A string is a sequence of character data.
- **Number:** A number is an integer that can be used in mathematical operations in code.

- **List:** A list is a container that is used for storing multiple data types at the same time.
- **Dictionary:** A dictionary uses key/value pairs to match keys to values.
- **Boolean:** A Boolean is a (AND, OR, NOT) operator that is used in code.
- **File:** A file in an object used by Python.

Before we look at strings, let's look at using the **print** statement. The **print** statement prints a specified message (which can be a string or any other object) to the screen. You can use the **print** command or type the variable name in the interpreter. The object is converted to a string before being written to the screen.

When you work with strings, you can use a few functions to modify or verify data within the string. These functions are known as *methods*.

Example 12.7 demonstrates the use of the **replace** method to change the **:** symbol to a period (**.**). Then, it uses the **startswith** method to verify that the IP address starts with the value **10**. It then uses the **format** method and the **{}** symbols to insert the value **100** in the IP address. Finally, the **split** method is used to create a list of the values, separated by the **.** character.

EXAMPLE 12.7 **String Methods**

```
>>> hostname = 'switch1'
>>> hostname.upper ()
'SWITCH1'
>>>
>>> macaddr = '52:54:00:1b:33:81'
>>> macaddr.replace(':', '.')
'52.54.00.1b.33.81'
>>>
>>> ipaddr = '10.1.1.1'
>>> ipaddr.startswith('10')
True
>>>
>>> ipaddr = '10.{}.1.1'
>>> ipaddr.format('100')
'10.100.1.1'
>>>
>>> ipaddr = '10.2.4.1'
>>> ipaddr.split('.')
['10', '2', '4', '1']
```

As noted earlier, a Boolean is an (AND, OR, NOT) operator that is used in code. The following Booleans are used in Python:

- **AND:** All values must match for the result to be True.
- **OR:** Any value must match for the result to be True.
- **NOT:** This Boolean takes the inverse.

Table 12.2 shows these Boolean (AND, OR, NOT) operators and their results.

TABLE 12.2 **Boolean Operators and Results**

| **AND** | **Result** |
|---|---|
| True and False | False |
| True and True | True |
| False and False | False |
| **OR** | **Result** |
| True or False | True |
| True or True | True |
| False or False | False |
| **NOT** | **Result** |
| not True | False |
| not False | True |

Example 12.8 demonstrates the results of Boolean AND, OR, and NOT operators.

EXAMPLE 12.8 **Using Boolean AND, OR, NOT Operators**

```
>>> True and False
False
>>>
>>> True or False
True
>>>
>>> not (True or False)
False
>>>
```

Objects and expressions in Python always evaluate to **True** or **False**, even for variables. This means that a variable such as **hostname = switch1** can specify

**if hostname** in a condition. Once there is a value assigned, and it is not null, the condition will evaluate to **True**; otherwise, it will be **False**. When working in Python, a conditional statement must end in an indentation, and it must end with a colon. The block of line indented after the colon (:) is executed whenever the condition is TRUE. The colon (:) is important since it separates the condition from the statements that are to be executed after the evaluation of the condition. The three conditional statements are **if**, **elif**, and **else**:

- **if:** The **if** statement specifies a conditional to be evaluated.
- **elif:** The **elif** optional conditional statement can be used numerous times to check for multiple expressions of **True**.
- **else:** The **else** conditional is the statement that contains the code to execute an action.

Example 12.9 demonstrates the use of the if, elif, and else conditional statements.

EXAMPLE 12.9 **Conditional Statements**

```
>>> switch = 'catalyst 9000'
>>>
>>> if 'catalyst' in switch:
...     switch_type = 'catalyst'
... elif 'nexus' in switch:
...     switch_type = 'nexus'
... else:
...     switch_type = 'unknown'
...
>>>
>>> print switch_type
catalyst
>>>
```

# Python Requirements

**ExamAlert**

For the ENCOR exam, you should remember the key points you need to adhere to when putting together Python scripts.

Because Python is an interpreted programming language, a Python file does not need to be compiled into a form that the machine understands before it is run. Python can interpret the file. Some key points that you should adhere to when putting together Python scripts are as follows:

- A Python filename ends with **.py** and is executed using the command **python** *filename*.**py**.
- The code in a Python file is the same as the code shown in the Python interactive interpreter.
- It is a good practice to include a shebang (discussed next) on the first line of a Python script.
- It is good practice to define an entry point.

It is generally recommended to use a shebang on the first line of a Python script. Although this is not mandatory, it helps enforce which Python version is used by the shell environment when executing a script. A shebang character sequence is a special character in a script file that is denoted by **# !**. It is always used in the first line of any file, and it helps in specifying the type of program that should be called to run the entire script file. A shebang, when used, is the only line that can start with a **#** that is not a comment. (For inline comments anywhere else in a script, you use **#** .) For example, if you use the shebang **# !/usr/bin/env python** when running a script, it will run the version of Python configured in the environment.

It is also recommended that you define an entry point for a Python script. With an entry point, you explicitly define where the code begins to get executed when you execute a Python file as a standalone script/program. Defining an entry point is optional. However, if an entry point is not defined, the code gets executed from the top down when the Python file is run or when it is being used as a Python module.

# Parsing Python Output to JSON

Let's look at how to parse Cisco IOS/Python output to JavaScript Object Notation (JSON) from the Python interpreter. JSON is a lightweight data interchange format that was inspired by JavaScript object literal syntax. The output of Cisco IOS is unstructured. For scripting purposes, you need to make sense of the output by making it structured to a format like JSON. One way of making the unstructured Cisco IOS data structured and more useful is to parse it to JSON format.

To get started working with JSON data, you can use Python's built-in module **json** to import this module by using the command **import json**. To make it easier to work with the data output, you can convert it into a Python dictionary and use lists. Breaking the JSON data into dictionary and list format makes it easier to access the item in that data structure.

A few Python methods are used to load JSON data:

- **load():** This method loads data from a JSON file into a Python dictionary.
- **loads():** This method loads data from a JSON variable into a Python dictionary.
- **dump():** This method loads data from a Python dictionary to a JSON file.
- **dumps():** This method loads data from a Python dictionary to a JSON variable.

# Exception Handling

Normally when an error or exception occurs, Python stops executing the code and generates an error message. If the variable does not exist, the program stops and does not move on to the next line. Python has a list of built-in exceptions:

- **try/except:** A **try** block lets you test a block of code for errors. The **except** block lets you handle an error that occurs in a **try** block.
- **else:** The **else** keyword is used to define a block of code that should be executed if there are no errors.
- **finally:** A **finally** block lets you execute code regardless of the rest of the **try** and **except** blocks.
- **raise:** The **raise** keyword is used to raise an exception.

You can also define multiple exceptions if you want to execute a special block of code for a special kind of error. You can, for example, define the type of error to raise and the text to print.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. True or false: Python is well suited for use in network troubleshooting and for event-based actions.
   - ❍ **A.** True
   - ❍ **B.** False

2. Which of the following data types uses key/value pairs to match keys to values?
   - ❍ **A.** String
   - ❍ **B.** Number
   - ❍ **C.** List
   - ❍ **D.** Dictionary

## Answers

1. **A** is correct. Python is used for network troubleshooting, general use, event-based actions, and other DevOps operations.
2. **D** is correct. Dictionaries use key/value pairs to match keys to values.

# Review Questions

1. True or false: Native, or onboard, Python refers to Python that is installed natively on a switching platform.
   - ❍ **A.** True
   - ❍ **B.** False

2. Which of the following is not a highlight of Python 3.x?
   - ❍ **A.** Supported on most newer platforms
   - ❍ **B.** Easier to learn
   - ❍ **C.** Can be used with virtualenv
   - ❍ **D.** Provides better library support

3. Which of the following commands is used for activating Guest Shell to run the Python interpreter on the Cisco IOS XE platform?
   - ❍ **A. iox**
   - ❍ **B. guestshell run**
   - ❍ **C. guestshell enable**
   - ❍ **D. guestshell run python**

4. True or false: When handling exceptions in Python, the **else** keyword defines a block of code that should be executed if no error is found.
   - ❍ **A.** True
   - ❍ **B.** False

# Answers to Review Questions

1. **A** is correct. Native, or onboard, Python refers to Python that is installed natively on a switching platform such as the Cisco Catalyst 9300.
2. **D** is correct. Python 2.x provides better library support than Python 3.x.
3. **C** is correct. Guest Shell is enabled using the command **guestshell enable**.
4. **A** is correct. The **else** keyword defines a block of code that should be executed if there are no errors.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers building JSON files.

CHAPTER 13

# Building JSON Files

**This chapter covers the following official ENCOR 350-401 exam objective:**

- 6.2 Construct valid JSON encoded file

Implementing network automation assists network administrators in managing complex networks. There are a number of ways to implement network automation. This chapter and the following few chapters look at using network-centric programming. This chapter looks at two common data formats that are used with application programming interfaces (APIs): Extensible Markup Language (XML) and JavaScript Object Notation (JSON).

**This chapter covers the following technology topic:**

- Data Formats (XML and JSON)

# Data Formats (XML and JSON)

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. What are the three protocols of the model-driven programmability stack?
2. Which data format is easier for humans to write: XML or JSON?

### Answers

1. NETCONF, RESTCONF, and gRPC
2. JSON is easier and faster for humans to write than XML.

With model-driven architectures, the software maintains a complete, explicit representation of the administrative and operational state of the system (the model) and performs actions only as side effects of mutations of the model entities. The model-driven programmability stack of Cisco network devices allows you to automate the configuration and control of the devices. Data models are written in a standard, industry-defined language and deliver a programmatic and standards-based method of writing configurations to network devices. Using a model replaces the process of manually configuring network devices. When configuring network devices, using a CLI may be human-friendly, but automating the configuration using data models results in better scalability.

The key to understanding data-encoding formats like XML and JSON is to have a solid understanding of where they fall within the model-driven programmability stack. This chapter and several others examine the different components of the model-driven programmability stack. You need to have good knowledge of these components in order to pass the ENCOR exam.

### ExamAlert

For the ENCOR exam, make sure you understand the components of the model-driven programmability stack as well as the data-encoding formats.

The following essential components make up the model-driven programmability stack:

- **Data models:** Data models are the foundations of APIs. They define the syntax and semantics and constraints involved in working with APIs.
- **Transport:** A model-driven API supports one or more transport methods (such as SSH, TLS, and HTTP/HTTPS).
- **Encoding:** Model-driven APIs support a number of encoding formats, including XML and JSON, which are the focus of this chapter. They can also include language-neutral/platform-neutral encoding, such as Google protocol buffers.
- **Protocols:** Model-driven APIs support several options for protocols, including the three core protocols NETCONF, RESTCONF, and gRPC.

XML and JSON are the two main data-encoding formats used in APIs. Let's look at XML before getting into JSON.

# Extensible Markup Language (XML)

XML, which is one of the most widely used text-based formats for describing data, was defined and refined by the World Wide Web Consortium (W3C). It is a commonly used format for sharing data between programs, computers, and people or between computers and people both locally and across networks.

The following are some of the main characteristics of XML:

- XML documents use a "self-describing" (that is, human readable) and simple syntax.
- The basic XML building block is an element that is defined by a tag pair that consists of a beginning tag and an ending tag.
- Elements can be nested within other elements.
- The outermost element is considered the root element.
- Each element can contain attributes.

XML is quite similar to HTML. However, it is not a replacement for HTML. XML is designed to transport and store data, with a focus on what the data is. On the other hand, HTML is designed to display data, focusing on how the data looks. Or, in other words, whereas HTML is used for displaying of information, XML focuses on carrying information. HTML uses predefined tags;

that is, HTML documents can only use tags that the HTML standard defines. In contrast, XML allows the author to define tags and document structure.

Often, computer systems and databases contain data in incompatible formats. XML stores data in plaintext format. Storing data in this format allows for a software- and hardware-independent method of storing data. This also makes it easier to create data that various applications can share. Getting data exchanged between incompatible systems can be both time-consuming and challenging. Using XML to exchange data significantly reduces this complexity because various incompatible applications can read the data.

As mentioned earlier, the syntax for XML is self-describing and simple. Let us take a closer look:

- The first line of an XML file is the XML declaration, which defines the XML version (for example, 1.0) and the encoding used (for example, ISO-8859-1, which is the Latin-1, Western European character set).
- The next line describes the root element of the document. The root element is a requirement and serves as the "parent" of all other elements.
- The next four lines describe four child elements of the root (to, from, heading, and body).
- The last line of an XML file defines the end-of-the-root element.

Together, the elements in an XML document form a document tree. The tree starts at the root and branches to the bottom level of the tree. All elements in an XML document can have sub-elements (that is, child elements).

Example 13.1 shows a snippet of an XML document.

EXAMPLE 13.1 **XML Snippet**

```
<?xml version="1.0" encoding="ISO-8859-1"?>

<switches>
  <switch name="switch01">
    <hostname>switch01</hostname>
    <mgmt-interface>mgmt0</mgmt-interface>
    <mgmt-ip>10.10.10.1</mgmt-ip>
    <mgmt-mask>255.255.255.0</mgmt-mask>
    <mgmt-gw>10.10.10.254</mgmt-gw>
  </switch>
</switches>
```

In this example, **<switches>** is the root element and carries the tag **switches**. The root element includes the element **<switch>**, which has the attribute **name**. As you can see, it is simple to build a hierarchy in an XML document by nesting additional elements inside an element. You can also see that each element has a beginning tag and an ending tag.

# JavaScript Object Notation (JSON)

Although XML is very flexible and can be used to represent complex data, using XML can at times be complicated. JSON was introduced as an alternative to XML, and it is the focus of this section.

JSON is a lightweight format that is used for data interchange and was first defined in RFC 4627 by Douglas Crockford. As you saw earlier in this chapter, web services historically used XML as the primary data format for transmitting data. Since its introduction, however, JSON has served as an alternative to XML and is often the preferred format because it is more lightweight.

The following are some of the main highlights of the JSON format:

- It describes data in plaintext and is human readable.
- It is language independent.
- It allows nested elements to provide hierarchy.
- The elements can be objects, arrays, or key/value pairs.

**ExamAlert**

For the ENCOR exam, make sure you fully understand the construction of JSON files. It is important to understand that objects are key/value pairs separated by colons and to understand how to use curly braces and arrays.

Like XML, JSON is plaintext and self-describing; in fact, it is even more human readable than XML. JSON is hierarchical and contains values within values. However, unlike XML, JSON has no end tags. It is generally shorter and quicker to read and write and has less overhead than XML. You can quickly parse JSON by using the built-in JavaScript function **eval()**. JSON uses arrays and has no reserved words.

JSON stores information using key/value pairs. A JSON file includes the following components:

- Objects, which start with **{** and end with **}**
- Comma-separated objects
- Double quotes, which enclose names and strings
- Lists, which start with **[** and end with **]**

In JSON, an object is framed using the symbol **{** at the beginning and the symbol **}** at the end. All of the key/value pairs are contained within this framing. When a JSON object comprises more entries than a single key/value pair, the entries are separated by commas. The comma separator is important, and if it is absent, errors are triggered, and the data isn't processed.

A key is always framed using quotation marks. However, the value is not always framed with quotation marks. It depends on the type of the value.

Now that we have looked at the structure of JSON files, let's look at its data types:

- **String:** Text that is enclosed in quotes
- **Number:** A positive or negative integer or a floating-point number
- **Object:** A key/value pair enclosed in curly braces
- **Array:** A collection of one or more objects
- **Boolean:** A value of either true or false with no quotes
- **Null:** The absence of data for a key/value pair

Example 13.2 shows a snippet of JSON formatting.

EXAMPLE 13.2 **JSON Snippet**

```
{
  "switches": [
    {
      "name": "switch01",
      "hostname": "switch01",
      "mgmt-interface": "mgmt0",
      "mgmt-ip": "10.10.10.1",
      "mgmt-mask": "255.255.255.0",
      "mgmt-gw": "10.10.10.254"
    }
  ]
}
```

A JSON document starts and ends with curly braces, meaning it holds objects. In Example 13.2, the name of the top-level object is **switches**, and it includes an array of switches. The array is enclosed in square brackets in the JSON document. The array consists of multiple JSON objects, one per switch. The **switch** JSON object contains multiple key/pair values that define the data.

## XML and JSON Comparison

Now that you have seen snippets and characteristics of both XML and JSON, let us look at the main features and highlights of both:

- JSON is more straightforward and compact than XML.
- XML is more verbose than JSON, but JSON is faster for humans to write.
- Parsing XML is difficult because of its complexity.
- JSON is better suited than XML for object-oriented systems.
- JSON is not as extensible as XML.
- JSON is better than XML for simple data exchange.
- JSON has less syntax than XML and no semantics.
- XML describes structured data that does not include arrays, whereas JSON includes arrays.
- The JavaScript **eval()** method parses JSON and returns the described object.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. In the model-driven programmability stack, which component is responsible for defining syntax and semantics?
   - ❍ **A.** Encoding
   - ❍ **B.** Transport
   - ❍ **C.** Data models
   - ❍ **D.** Protocols

2. True or false: JSON is simpler and more compact than XML and is better suited for object-oriented systems.
   - ❍ **A.** True
   - ❍ **B.** False

# Answers

1. **C** is correct. Data models are the foundations of APIs. They define the syntax and semantics and constraints involved in working with APIs.
2. **A** is correct. JSON is simpler and more compact than XML, is easier to parse due to its simplicity, and is more suited for object-oriented systems.

# Review Questions

1. The first line of an XML document indicates which of the following?
   - ❍ **A.** Heading
   - ❍ **B.** Root element
   - ❍ **C.** Child element
   - ❍ **D.** XML declaration

2. True or false: A JSON document contains an end tag.
   - ❍ **A.** True
   - ❍ **B.** False

3. What JSON construct is used to separate key/value pairs?
   - ❍ **A.** Colon
   - ❍ **B.** Comma
   - ❍ **C.** String
   - ❍ **D.** Array

4. True or false: With JSON, the absence of a comma separator triggers errors.
   - ❍ **A.** True
   - ❍ **B.** False

# Answers to Review Questions

1. **D** is correct. The first line is the XML declaration. It defines the XML version (for example, 1.0) and the encoding used (for example, ISO-8859-1, which is the Latin-1, Western European character set).
2. **B** is correct. JSON is hierarchical and contains values within values. However, unlike XML, JSON has no end tags.
3. **B** is correct. A comma is used to separate every key/value pair in JSON.
4. **A** is correct. The absence of a comma separator in JSON triggers an error in the application.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*
- *Network Programmability and Automation Fundamentals*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers YANG data modeling.

CHAPTER 14

# YANG Data Modeling

**This chapter covers the following official ENCOR 350-401 exam objective:**

- 6.3 Describe the high-level principles and benefits of a data modeling language, such as YANG

A data model describes what can be configured on a device, what can be monitored on the device, and all administrative actions that can be executed. Data models are powerful in the sense that they allow you to create uniform ways of describing data, which can be helpful across vendors' platforms. This chapter looks at Yet Another Next Generation (YANG) data models. YANG, which is used as an alternative to Simple Network Management Protocol (SNMP) management information bases, uses a tree structure to describe data.

**This chapter covers the following technology topic:**

- YANG Data Modeling

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. What protocols does YANG use for communication?
2. Which YANG node type contains a sequence of leaf nodes?

### Answers

1. NETCONF, RESTCONF, and gRPC
2. Leaf-list node

# YANG Data Modeling

Several data models and tools are commonly leveraged in a programmatic approach:

- YANG modeling language
- Network Configuration Protocol (NETCONF)
- Representational State Transfer Configuration Protocol (RESTCONF)

This chapter focuses on YANG, and Chapter 32, "NETCONF and RESTCONF," discusses the other two data models.

Traditionally, Cisco network devices were managed using the CLI for configuration (configuration commands) and operation (**show** commands). Network management was typically handled using Simple Network Management Protocol (SNMP). However, these tools present several restrictions. The Cisco CLI is proprietary and requires human intervention for understanding and interpreting the text-based specifications. SNMP does not distinguish between configuration and operational data.

The solution to these shortcomings lies in adapting a programmatic and standards-based method of writing configurations to any network device, thus replacing the manual configuration process. For example, network devices running Cisco IOS XE and NX-OS can be configured using automation based on data models. A data model is developed in a standard-based,

industry-defined language that can identify the configuration and state information of the network. Let us briefly look at the benefit of using data models before getting into YANG.

These are some of the benefits of using data models:

- They provide a common model for configuration and operational state data, and they can perform NETCONF actions.
- They use protocols to communicate with a network device to get, manipulate, and delete configurations in a network.
- They automate configuration and operation of multiple network devices across the network.

YANG is a standard-based data modeling language initially defined in RFC 6020. YANG is maintained by the NETMOD working group within the Internet Engineering Task Force (IETF). It is used to create device configuration requests or for making requests for operational data. YANG models the hierarchical organization of data in a tree structure, in which each node has a name and a value, or a set of child nodes. YANG provides clear and concise descriptions of nodes and the interaction between nodes. In addition, YANG provides reusable structures that can be used within and between YANG models.

ExamAlert

For the ENCOR exam, you need to know that there are both standards-based and vendor-specific YANG models.

# Different YANG Models

There are a number of YANG models, including standard-based as well as vendor-specific models:

- **Industry standard-based (common) models:** These models apply to all vendors and come from various working groups, including the IETF and the Open Config working group. The purpose of this group is to create vendor- and platform-independent models with core features that are relevant across different vendors. For example, a request to enable or shut down an Ethernet interface should be identical for Cisco and non-Cisco devices.

- **Cisco common models:** Because a number of features are common across different Cisco devices, there are Cisco native models as well as native models for each Cisco OS.
- **Cisco platform-specific models:** These models facilitate the configuration or collection of operational data associated with a Cisco platform or hardware-specific features. These models ensure that features that are mapped to a particular platform are still model driven, meaning that APIs can still be used for those features.

A YANG module defines a single data model. However, a module can reference definitions in other modules and submodules, using the following statements:

- **import:** Imports external modules
- **include:** Includes one or more submodules
- **augment:** Provides augmentations to another module and defines the placement of new nodes in the data model hierarchy
- **when:** Defines conditions under which new nodes are valid
- **prefix:** References definitions in an imported module

> **ExamAlert**
>
> For the ENCOR exam, ensure that you understand the communication protocols that YANG uses.

Communication protocols establish connections between network devices and clients. These protocols help a client consume the YANG data models, which, in turn, automate and program network operations. YANG uses these protocols for communication:

- Network Configuration Protocol (NETCONF)
- Representational State Transfer Configuration Protocol (RESTCONF)
- RPC framework (gRPC) by Google

YANG can be used with NETCONF to provide automated and programmable network operations. NETCONF is an XML-based protocol that client applications use to request information from and make configuration changes to

network devices. YANG is primarily used to model the configuration and state data that is used by NETCONF operations. In other words, YANG can use NETCONF as the medium to move configuration and operational data from a network device to an application. (Both NETCONF and RESTCONF are covered in Chapter 32.)

gRPC is an open-source remote procedure call (RPC) framework based on Protocol Buffers, an open-source binary serialization protocol. gRPC offers a flexible, efficient, and automated mechanism for serializing structured data, like XML, but it is smaller and simpler to use than XML.

YANG data models can be represented in a hierarchical, tree-based structure with nodes. This representation makes the models easy to understand. Each feature has a defined YANG model, which is synthesized from schemas. A model in a tree format includes the following:

- Top-level nodes and their subtrees
- Subtrees that augment nodes in other YANG models
- Custom RPCs

In YANG, each node has a name and is one of four types that either defines a value or contains a set of child nodes:

- **leaf node:** A leaf node contains a single value of a specific type.
- **leaf-list node:** A leaf-list node contains a sequence of leaf nodes.
- **list node:** A list node contains a sequence of leaf-list entries, each of which is uniquely identified by one or more key leafs.
- **container node:** A container node contains a group of related nodes that have only child nodes, which can be any of the four node types.

The following example looks at the structure of a Cisco Discovery Protocol (CDP) data model on an IOS XE device. There are a number of YANG modules. In this case, we will look at a CDP module structure in tree format, and then we will look at how it is expressed in YANG. Cisco CDP configuration has an inherent augmented model (interface configuration).

Example 14.1 shows the data model for a CDP interface manager in tree structure. The augmentation indicates that CDP can be configured both at the global configuration level and at the interface configuration level.

EXAMPLE 14.1 **CDP Interface Manager in Tree Structure**

```
module: Cisco-IOS-XE-cdp
   +--rw cdp
      +--rw timer?                uint32
      +--rw advertise-v1-only?    empty
      +--rw enable?               boolean
      +--rw hold-time?            uint32
      +--rw log-adjacency?        empty
augment /a1:interface-configurations/a1:interface-configuration:
   +--rw cdp
      +--rw enable?   empty
```

Example 14.2 shows the augmentation expressed in the CDP YANG model. It shows how a CDP tree structure is expressed in the tree/subtrees format for the interface-level CDP configuration.

EXAMPLE 14.2 **CDP YANG Model**

```
augment "/a1:interface-configurations/a1:interface-configuration" {
    container cdp {
      description "Interface specific CDP configuration";
      leaf enable {
        type empty;
        description "Enable or disable CDP on an interface";
      }
    }
    description
      "This augment extends the configuration data of
      'Cisco-IOS-XE-cdp'";
  }
```

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which of the following is not a common data model that is used in a network programmatic approach?
   - ❍ **A.** YANG
   - ❍ **B.** SNMP
   - ❍ **C.** NETCONF
   - ❍ **D.** RESTCONF

2. Which YANG model was designed to create vendor and platform-independent models?

   - ❍ **A.** Common
   - ❍ **B.** Cisco common
   - ❍ **C.** Cisco platform-specific
   - ❍ **D.** IEEE

# Answers

1. **B** is correct. SNMP is not a common data model in a network programmatic approach as it does not distinguish between configuration and operational data.
2. **A** is correct. Industry standard-based (common) models apply to all vendors and come from various working groups, including the IETF and the Open Config working groups. The purpose of this group is to create vendor- and platform-independent models with core features that are relevant across different vendors.

# Review Questions

1. Which YANG node type contains a sequence of leaf-list entries?
   - ❍ **A.** Leaf node
   - ❍ **B.** Leaf-list node
   - ❍ **C.** List node
   - ❍ **D.** Container node

2. True or false: YANG can use NETCONF as the medium to move configuration and operational data.
   - ❍ **A.** True
   - ❍ **B.** False

# Answers to Review Questions

1. **C** is correct. A list node contains a sequence of leaf-list entries, each of which is uniquely identified by one or more key leafs.
2. **A** is correct. YANG can use NETCONF as the medium to move configuration and operational data from a network device to an application.

# Further Reading

- ▶ *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*
- ▶ *Network Programmability and Automation Fundamentals*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers DNA Center and vManage APIs.

CHAPTER 15

# DNA Center and vManage APIs

**This chapter covers the following official ENCOR 350-401 exam objective:**

- 6.4 Describe APIs for the Cisco DNA Center and vManage

This chapter examines how the Cisco DNA Center and Cisco vManage make use of REST APIs for communication. Chapter 16, "Interpreting REST API Codes," follows up by discussing REST API codes.

**This chapter covers the following technology topic:**

- APIs for Cisco DNA Center and vManage

## Cram Saver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. What facilitates the building of bidirectional interfaces to allow the exchange of contextual information between Cisco DNA Center and external third-party systems?
2. Which site management API manages software images and the update repository for network devices?

### Answers

1. REST-based integration adapter APIs
2. Software Image Management (SWIM)

# APIs for Cisco DNA Center and vManage

DNA Center is covered exclusively in Chapter 23, "SD-Access" and vManage in Chapter 22, "SD-WAN." This chapter provides a brief overview of Cisco DNA Center and Cisco vManage.

# DNA Center API Integrations

Cisco DNA Center is at the heart of Cisco's intent-based network architecture. It supports intent-based networking by expressing the business intent for network use cases. Cisco DNA Center features allow for the creation of value-added applications to leverage the full capabilities of DNA Center.

The main goal of the Cisco DNA Center platform is to streamline end-to-end IT processes across the value chain. Cisco DNA Center achieves this through integration with various ecosystem domains, such as IT Service Management (ITSM), IP Address Management (IPAM), and reporting.

Figure 15.1 shows the Cisco SD-WAN architecture, which provides support for third-party integration.

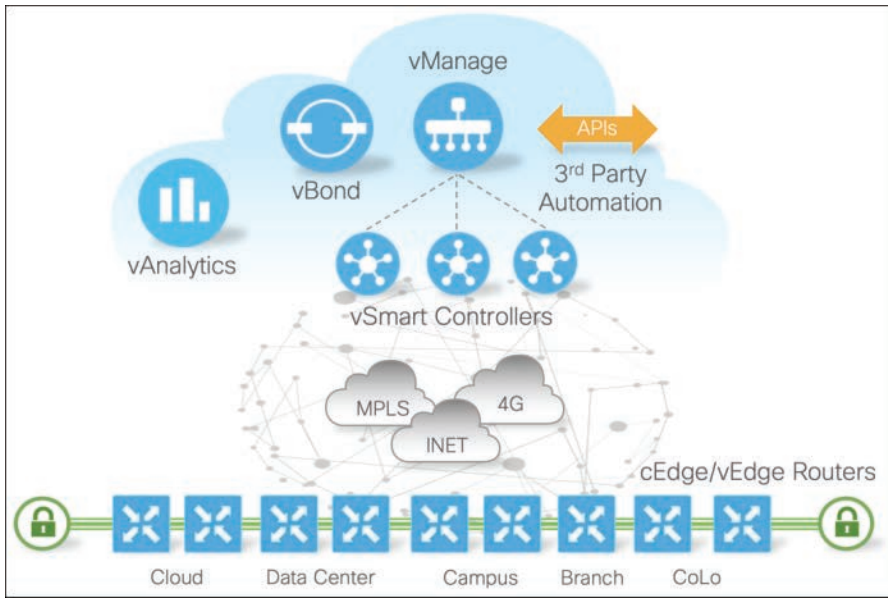



FIGURE 15.1 **Cisco SD-WAN Architecture**

By using REST-based integration adapter APIs, you can build bidirectional interfaces to allow the exchange of contextual information between Cisco DNA Center and external third-party systems. The APIs provide the capability to publish the network data, events, and notifications to the external systems. The APIs also provide the ability to consume information in Cisco DNA Center from the connected systems.

You can use the Cisco DNA Center Intent API, integration flows, events and notification services, and the optional Cisco DNA Center multivendor SDK to improve the overall network experience. This may involve optimizing end-to-end IT processes, reducing total cost of ownership (TCO), and creating value-added networks.

**ExamAlert**

Before taking the ENCOR exam, make sure you understand the HTTP operations that the RESTful DNA Center API allows you to use. Also make sure you understand, at a high level, how the Know Your Network request paths are organized.

The RESTful DNA Center API allows you to use HTTP objects (GET, PUT, POST, and DELETE) and JSON syntax to discover and control networks. You

can use the Know Your Network REST request paths to retrieve details about clients, sites, topology, and devices.

The Developer Toolkit > APIs page organizes Know Your Network request paths into the following subdomains:

- **Sites:** A site is a hierarchal collection of other sites and buildings. You can create and manage sites, assign devices to sites, or obtain site information, site count, and site membership.
- **Topology:** The topology retrieves network health information, site, physical network, Layer 2, Layer 3, and VLAN information.
- **Devices:** You can create, manage, and retrieve detailed information about devices based on a wide range of attributes, such as timestamp, MAC address, UUID, name or device name, functional capabilities, interfaces, device configurations, certificate validation status, values of specified fields, modules, and VLAN data associated with specific interfaces.
- **Clients:** You can obtain (GET) detailed client and client health information.
- **Users:** You can obtain (GET) detailed information about a user, given an identifying network user ID or MAC address.
- **Issues:** You can obtain (GET) detailed information and recommended mitigation for an issue, given an identifying issue ID or MAC address.

Site management enables you to provision networks with zero-touch provisioning. Site management APIs assist with activation and distribution of software images. The following are examples of the site management APIs:

- **Site Design:** These methods are used to create and obtain information about provisioned Network Functions Virtualization (NFV) devices.
- **Network Settings:** These methods can get device credentials, global pool information, and service provider details.
- **Software Image Management (SWIM):** SWIM manages the software image and update repository for network devices. Software images can be stored in Cisco DNA Center or imported from a designated URL.
- **Device Onboarding (PnP):** These methods support the management of device onboarding projects, settings, workflows, virtual accounts, and PnP-managed devices. They allow for the zero-touch deployment of Cisco enterprise network routers, switches, wireless controllers, and wireless access points.

- **Configuration Templates:** Configuration templates with the Template Programmer/Editor are centralized CLI-management tools that facilitate the design and provisioning of workflows in Cisco DNA Center. CLI templates can be grouped into projects. These methods enable management of CLI templates.

The operational tools support access to CLI keywords, allowing you to discover network devices, configure network settings, and trace paths through the network. The following are examples of operational tools:

- **Command Runner:** These methods support the retrieval of CLI keywords and enable the execution of read-only commands on a target network device.
- **Network Discovery:** Discovery is the process of scanning a target network to add existing network devices to the device inventory. Discovery is conducted using SNMP or other protocols.
- **Path Trace:** These methods provide flow analysis between two endpoints on a network. You can initiate new path trace analyses and read prior path traces.
- **File:** File services are used to list file namespaces and to download specific files from DNA Center.
- **Task:** Tasks are Cisco DNA Center activities that are initiated for asynchronous execution via an API request. These are GET methods that include status information about a single task, indicating completion or progress.
- **Tags:** A tag is a named set of attributes associated with a member. Tag services provide the means to create, discover, update the membership of, and remove tags.

Connectivity methods provide the means necessary for configuring and managing fabric wired and non-fabric wireless networks:

- **Fabric wired:** These methods are used to manage fabric wired devices, including creation, update, and deletion of edge, border, user devices, and authentication profiles.
- **Non-fabric wireless:** These methods are used to manage and provision non-fabric wireless devices, including enterprise SSIDs, wireless profiles, RF profiles, and access points.

Cisco DNA Center allows you to create application policies that reflect your organization's business intent and translate them into network-specific and device-specific configurations required by the different types, makes, models, operating systems, roles, and resource constraints of your network devices. Application policy methods support the creation, update, and management of applications and application sets.

Cisco DNA Center enables you to receive custom notifications when specific events are triggered and enables third-party systems to take business actions in response to certain events. Continuing with DNA Center API, you have the following integrations to enhance the overall network experience.

- **Multivendor support (southbound):** Cisco DNA Center enables you to manage non-Cisco devices through the use of a software development kit (SDK) that can create device packages for third-party devices. Device packages enable Cisco DNA Center to communicate to third-party devices by mapping Cisco DNA Center features to their southbound protocols. Integration with third-party components provides an integrated view of the network that is consistent with the DNA Center abstraction.
- **Events and notifications (eastbound):** Cisco DNA Center can establish a notification handler to operate on specific events that are triggered. This mechanism enables external systems to take actions in response to certain events. For example, for a network device that is out of compliance, a custom application can execute a software upgrade in reaction to a notification about its noncompliance.
- **Integration API (westbound):** Integration capabilities are part of westbound interfaces. To scale operation in modern data centers, you need intelligent end-to-end workflows created with open APIs.

Cisco DNA Center provides mechanisms for integrating Cisco DNA Assurance workflows with third-party IT service management (ITSM) solutions.

# vManage API Integrations

Let us now look at how Cisco Software-Defined WAN (SD-WAN) uses APIs for communication and to implement policy changes within its fabric.

The Cisco SD-WAN solution using vManage for management offers an SD-WAN fabric with centralized management and security built in, creating a secure overlay WAN architecture across campus, branch, and data center and multi-cloud applications. The SD-WAN software solution runs on a range of SD-WAN routers across hardware, virtual, and cloud form factors.

The Cisco SD-WAN vManage API is a REST API interface used for controlling, configuring, and monitoring the Cisco devices in an overlay network. The API plays a crucial role for clients to consume the features provided by vManage.

These are some of the common use cases for the vManage API:

- Monitoring device status
- Configuring a device, such as attaching a template to a device
- Querying and aggregating device statistics

The vManage API can be categorized in the following categories:

- **Administrative and management APIs:** These APIs handle user, group, tenant management, software maintenance, backup and restore, and container management.
- **Alarm and event APIs:** These APIs handle alarm and event notification configuration as well as alarm, event, and audit log queries.
- **Configuration APIs:** These APIs handle template, device templates, device policies, device certificates, device actions, action status, device inventory, and so on.
- **Device real-time monitoring APIs:** These APIs handle real-time monitoring of devices, links, applications, systems, and so on.
- **Device state statistics bulk APIs:** These APIs handle device states, aggregated statistics, and bulk queries.
- **Troubleshooting and utility APIs:** These APIs handle troubleshooting of devices and systems.

Before taking a closer look at the REST response, let us review some of the steps you take to connect to the vManage REST API:

1. To issue a REST API call, place the call in the URL field: https://*vmanage-ip-address:port/api-call-url*.
2. If you need to make an API call to retrieve a list of all the network devices in the network, use the following: https://*vmanage-ip-address:port/dataservice/device*. This call returns a JSON object, and it may be large as it contains device information for all devices in the network.
3. Filter the results to get information for only a single device. For this, you need to add query string parameters. For example, to limit the GET response to a single device, you specify a particular system IP address

as follows: https://*vmanage-ip-address:port/dataservice/device?system-ip=192.168.1.254*. In this case, the JSON object returns the output for only that device.

4. At any point, access the API documentation by navigating to https://*vmanage-ip-address:port/apidocs*.

You can use the Postman development tool to make REST API calls to Cisco DNA Center or Cisco vManage infrastructure to authenticate and get a list of devices, device status, and interface statistics for all devices in the fabric. The following high-level steps are involved:

1. Set up authentication by using the Postman environment, where you can have a list of variables that can be used to switch between environments easily. (You simply need to change the vManage server IP address, username, password, and port.) The variables that you define for the environment can be reused throughout the API calls defined in the environment. The authentication call is a POST call. The API call is sent to the j_security_check resource. If authentication is successful, you should receive a 200 (OK) status message to indicate that authentication was successful.

2. Download post collectors—which are groups of API calls that define endpoints or resources available for a specific API—and add them to Postman. The collector and environment variables file can be downloaded from GitHub at https://github.com/CiscoDevNet/Postman-for-AlwaysOn-Cisco-SD-WAN. This public repository contains the Postman collection and environment files to interact with Cisco DevNet Always On Sandbox for SD-WAN 19.2 fabric (which is the most up-to-date version at this writing).

3. Once you are successfully authenticated, use a GET API from your collection. Some of the APIs you can experiment with to get some hands-on experiences are as follows:

   - POST Authentication
   - GET Fabric Devices
   - GET Devices Status
   - GET Device Counters
   - GET Interface Statistics

A common REST principle is that APIs should be self-descriptive and self-documenting. The resource collections and resources in the Cisco SD-WAN REST API are self-documenting, describing the data you need to make a call and the response from each call.

The vManage REST API library and documentation are bundled with and installed on the vManage web application software. To access the API documentation from a web browser, use the URL https://ip-address:port/apidocs.

Let us now look at the steps that are necessary to perform REST API operations on a vManage web server. Data transfer from a vManage web server using a utility such as Python goes through this process:

1. Establish a session to the vManage web server.
2. Issue the desired API call.

When you need to use a program or script to transfer data from a vManage web server or perform operations on the server, you first need to establish an HTTPS session to the vManage server. This involves sending a call to log in to the server with the following parameters:

- **The URL where request is sent:** https://*{vmanage-ip-address}*/j_security_check makes the login operation and security check on the vManage web server at the specified IP address.
- **Request method:** This specifies a Post request.
- **API call input:** The input is where you specify the application.
- **API call payload:** The payload contains the username and password in the format *j_username=username*&*j_password=password*.

Figure 15.2 shows the environment variables JSON file that contains all the details to authenticate to the Cisco DevNet SD-WAN Always On environment.

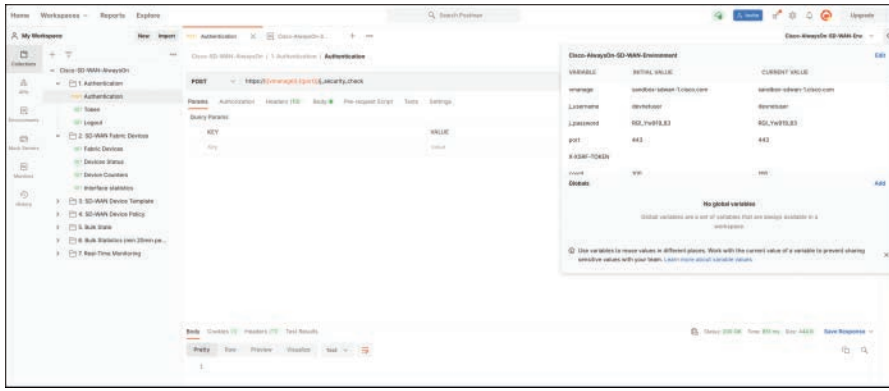


FIGURE 15.2 **Cisco DevNet SD-WAN Always On Environment JSON Details**

Figure 15.3 shows the output of the GET Fabric Devices API Call from vManage.

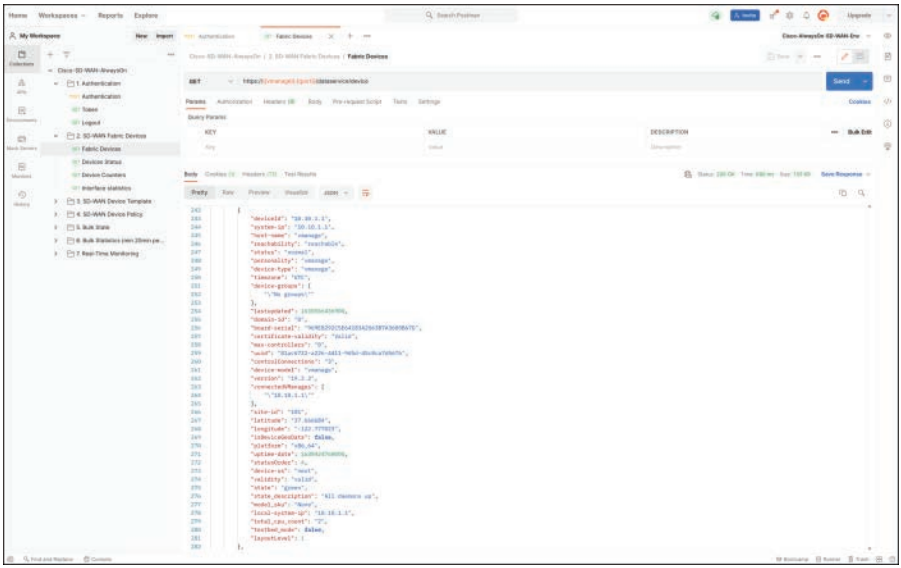


FIGURE 15.3 **GET Fabric Devices API Call**

Figure 15.4 shows the output of the GET Devices Status API Call from vManage.

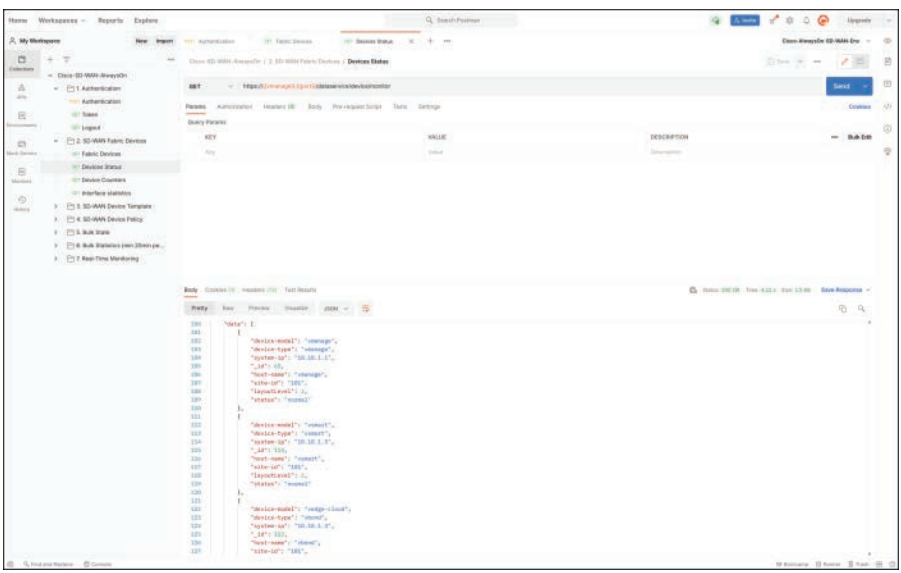


FIGURE 15.4 **GET Devices Status API Call**

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. True or false: Cisco DNA Center allows for the integration of various third-party ecosystem domains.
   - ❍ **A.** True
   - ❍ **B.** False

2. Which of the following is *not* a subdomain in the Know Your Network request path?
   - ❍ **A.** Devices
   - ❍ **B.** Clients
   - ❍ **C.** Sites
   - ❍ **D.** Environment

## Answers

1. **A** is correct. Cisco DNA Center allows integration with various ecosystem domains, such as IT Service Management (ITSM), IP Address Management (IPAM), and reporting.
2. **D** is correct. Environment is not one of the subdomains in the Know Your Network request path.

# Review Questions

1. True or false: Configuring a device, such as attaching a template to a device, is a common use case for Cisco vManage APIs.
   - ❍ **A.** True
   - ❍ **B.** False

2. True or false: The RESTful DNA Center API allows you to use HTTP operations like GET, PUT, POST, and DELETE to discover and control networks.
   - ❍ **A.** True
   - ❍ **B.** False

# Answers to Review Questions

1. **A** is correct. Configuring a device, such as attaching a template to a device, monitoring device status, and querying and aggregating device statistics, are common use cases for Cisco vManage APIs.
2. **A** is correct. The RESTful DNA Center API allows you to use HTTP verbs (GET, PUT, POST, and DELETE) and JSON syntax to discover and control networks.

# Further Reading

- ▶ *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*
- ▶ *Network Programmability and Automation Fundamentals*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers interpreting REST API codes.

CHAPTER 16

# Interpreting REST API Codes

**This chapter covers the following official ENCOR 350-401 exam objective:**

- 6.5 Interpret REST API response codes and results in payload using Cisco DNA Center and RESTCONF

This chapter covers the interpretation of REST API response codes. It looks at what happens when you log into a vManage web server to make requests to pull data from the vManage REST API. This is one of the shorter chapters in our automation review.

**This chapter covers the following technology topic:**

- Interpreting REST API Response Codes

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. What HTTP operation is used to send a username and password to Cisco DNA Center or Cisco vManage?
2. Which DNA Center HTTP code indicates that a resource was not found?

### Answers

1. POST
2. Status code 404

# Interpreting REST API Response Codes

As you saw in Chapter 15, "DNA Center and vManage APIs," vManage exposes the REST APIs. Because the DNA Center and vManage GUI is API driven, any action you trigger on the DNA Center or vManage GUI has an associated RESTful API call. The Cisco DNA Center REST API gives you a programmatic interface for controlling, configuring, and monitoring the Cisco SD-Access network devices that are part of an overlay network.

The Cisco DNA Center Intent API uses the GET, POST, PUT, and DELETE HTTP actions:

- **GET:** This action is used to retrieve data from the path that the URL specifies.
- **POST:** This action is used to write new data at the path that the URL specifies. It is used only to create new data on the server.
- **PUT:** This action is used to replace existing data at the URL path. This is not used to create new data but only to update existing data.
- **DELETE:** This action is used to remove existing data at the URL path.

ExamAlert

Before taking the ENCOR exam, make sure you understand the HTTP status codes in Cisco DNA Center and what they mean.

# HTTP Status Codes

When a web service processes an HTTP or REST request, it returns a standard HTTP status code to let you know whether the request succeeded. If an HTTP or REST request is successful, the web service returns a *2xx* response code. These are two of the common *2xx* response codes:

- **200 (OK):** The request was successful, and the response includes the result of the action that you requested. For example, if you create a user account, the status value 200 means that the account was created successfully.
- **202 (Accepted):** The server successfully accepted your request, but it needs to perform more actions to complete the request. For example, if Cisco DNA Center needs to interact with network switches to fulfill a request, it returns a 202 status code.

Status codes can also indicate that a request did not complete successfully. If an error was caused by a client request, the server returns a *4xx* status code. For example, status code 404 (Not Found) means that the resource was not found.

If an error was caused by a server error, the server returns a *5xx* status code. For example, the status code 503 (Service Unavailable) indicates that the server was unavailable to fulfill the request at the time the request was made.

Table 16.1 shows all of the Cisco DNA Center and vManage HTTP status codes used for troubleshooting.

TABLE 16.1 **Cisco vManage HTTP Status Codes**

| Status Code | Status Message | Meaning |
|---|---|---|
| 200 | OK | Success |
| 201 | Created | New resource was created |
| 400 | Bad request | Request was invalid |
| 401 | Unauthorized | Authentication was missing or incorrect |
| 403 | Forbidden | Request was understood but not allowed |
| 404 | Not Found | Resource was not found |
| 429 | Too Many Requests | Requests exceed rate limit |
| 500 | Internal Server Error | Problem with the server |
| 503 | Service Unavailable | Server is unable to complete request |

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. True or false? An error caused by a server typically prompts a 5*xx* message.
   - ❍ **A.** True
   - ❍ **B.** False

2. What HTTP status code indicates successful authentication to the Cisco vManage API or DNA Center?
   - ❍ **A.** 200
   - ❍ **B.** 201
   - ❍ **C.** 400
   - ❍ **D.** 401

# Answers

1. **A** is correct. If an error was caused by a server error, the server returns a 5*xx* status code.
2. **A** is correct. If authentication is successful, you should receive a status 200 (OK) message to indicate that redundant.

# Review Questions

1. True or false: HTTP status code 202 indicates that the server successfully accepted your request but needs more information.
   - ❍ **A.** True
   - ❍ **B.** False

2. True or false: HTTP status code 201 indicates that the server was unavailable to fulfill the request.
   - ❍ **A.** True
   - ❍ **B.** False

# Answers to Review Questions

1. **A** is correct. HTTP status code 202 means the server successfully accepted your request, but it needs to perform more actions to complete the request.
2. **B** is correct. HTTP status code 201 means that the resource was successfully created. HTTP statue code 503 indicates that the server is unable to complete the request.

# Further Reading

- ▶ *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*
- ▶ *Network Programmability and Automation Fundamentals*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers EEM applets.

*This page intentionally left blank*

CHAPTER 17

# EEM Applets

**This chapter covers the following official ENCOR 350-401 exam objective:**

- 6.6 Construct EEM applet to automate configuration, troubleshooting, or data collection

This chapter covers the operation of the Cisco Embedded Event Manager (EEM) built-in IOS tool as well as some common use cases for EEM. EEM is a unique subsystem in Cisco IOS software. It is powerful and flexible, allowing you to build software applets to automate tasks and customize the operation of the IOS software operating on devices. This chapter provides a brief overview of EEM and how to create EEM policies to illustrate common use cases: a policy that creates an applet to automatically enable a router interface if it goes down and a policy that creates a Tool Command Language (Tcl) script to ping multiple destinations.

**This chapter covers the following technology topic:**

- Embedded Event Manager (EEM)

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. What is the name of the EEM architecture component that is used to determine when an EEM event occurs?
2. Which EEM command is used to see the actions taking place on the CLI when the applet is running?

### Answers

1. Event detectors
2. **debug event manager action cli**

# Embedded Event Manager (EEM)

Embedded Event Manager (EEM) is a unique subsystem that is embedded within the Cisco IOS software. EEM is a powerful tool that automates tasks that you would otherwise typically have to execute manually from the command-line interface (CLI). EEM enables you to harness the intelligence within the Cisco IOS software to respond to real-time events, automate tasks, create customized commands, and take local automated action based on conditions detected by the Cisco IOS software. With EEM, you can create and run programs or scripts directly on a router or switch. The scripts created are referred to as *EEM policies* and can be programmed using either of the following:

- A CLI-based interface
- A scripting language called Tool Command Language (Tcl)

Although TCL is not specifically listed in the ENCOR exam objectives, using an EEM applet to call a TCL script is a powerful aspect of EEM. The focus on TCL in this chapter is limited to creating a basic TCL script and using an EEM applet to call a TCL script that is stored in flash memory.

EEM provides several features and benefits. It includes more than 20 event detectors that offer an extensive set of conditions that can be monitored and defined as event triggers. EEM is mostly product independent and available

across a wide range of Cisco routing and switching platform products. Cisco also introduces new event detectors and capabilities with each new version of EEM.

There are several use cases for EEM. For example, you might want to use it to react to an abnormal condition, such as forcing transit traffic over a more stable and error-free path upon detecting a high error rate on an interface. EEM can watch for the increased error rate and trigger a policy into action that could notify the network administrator and immediately take action to reroute traffic.

Another use case for EEM might be to collect detailed data upon detection of a specific failure condition. EEM can gather information needed to determine the root cause of a problem, and in doing so, it can reduce the mean time to repair and increase availability. EEM can also detect a specific syslog message and trigger a script to collect detailed data using a series of **show** commands. After automatically collecting the data, EEM can then save it to flash memory, send it to an external management system, or email it to a network administrator.

Another use case for EEM applets is to match CLI patterns to events. When certain commands are entered into the router using the CLI, an EEM event can be triggered within the applet. Then the configured actions can take place as a result of the CLI pattern being matched.

EEM puts a lot of control in a network administrator's hands as it provides control over what events to detect and what actions to take. It gives you flexibility in terms of when to use it, and you can configure it to only take actions that you configure it to take.

EEM scripts have two primary purposes:

- **Assist in troubleshooting an issue:** EEM scripts can assist in troubleshooting problems that occur intermittently. EEM allows you to automate the collection process with **show** command output and **debug** commands and enables you to capture data that would otherwise be hard to gather.
- **Assist with a temporary workaround:** EEM scripts can assist in cases where you need a temporary workaround while determining the root cause of an issue. For example, if you have a problem that is intermittent but that is fixed by an interface reset, you can use EEM scripts to trigger a reset as soon as the problem occurs so you can manually intervene to find the root cause of the problem.

To use EEM scripts, you must identify a trigger event that you can then use to trigger the script.

# EEM Architecture

The architecture of EEM consists of a series of event detectors, an EEM server, and interfaces to allow action routines (called policies) to be invoked. There are also internal application programming interfaces (APIs) for other Cisco IOS software subsystems to take advantage of the EEM subsystem.

**ExamAlert**

For the ENCOR exam, make sure you know the components that makes up the EEM architecture.

Let us briefly look at the main components that make up an EEM architecture:

- **EEM server:** The EEM server bridges the Cisco IOS subsystems used in the event detectors and the policies. Its primary purposes are to receive notifications from event detectors when an event of interest occurs, store the information about an event, publish events, register internal script directories, register Tcl scripts and applets, and process the actions taken by user-defined scripts.
- **Event detectors:** The event detectors in EEM are used to determine when an EEM event occurs. Event detectors are separate systems that provide an interface between the agent being monitored, like Simple Network Management Protocol (SNMP), and the EEM policies where an action can be implemented. The following are some examples of EEM event detectors:
  - **CLI event detector:** The CLI event detector screens CLI commands for a regular expression match. When there is a match, an event is published.
  - **Counter event detector:** The counter event detector publishes an event when a named counter crosses a specified threshold.
  - **Enhanced object tracking (EOT) event detector:** The EOT event detector publishes an event when the status of a tracked object changes. Each tracked object is identified by a unique number specified on the tracking CLI, and client processes use this number to track a specific object.
  - **IP service-level agreement (SLA) event detector:** The IP SLA event detector publishes an event when an IP SLA action is triggered.
  - **Routing event detector:** The routing event detector publishes an event when a route entry changes in the routing information base (RIB).

- **SNMP event detector:** The SNMP event detector allows a standard SNMP management information base (MIB) object to be monitored and an event to be generated when the object matches specified values or crosses stipulated thresholds.
- **Syslog event detector:** The syslog event detector allows you to screen syslog messages for regular expression pattern matches.
- **None event detector:** The none event detector publishes an event when the Cisco IOS **event manager run** CLI command manually executes an EEM policy.

# EEM Policies

As mentioned earlier in this chapter, EEM policies can be programmed in two ways:

- **Applet policies:** These policies can be created using an easy-to-use CLI.
- **Tcl policies:** These policies, which are defined using the Tcl programming language, can provide more flexible and extensive capabilities than applet policies.

**ExamAlert**

Before taking the ENCOR exam, you need to understand the high-level steps for setting up EEM as well at the more common actions that can be configured when an EEM applet is triggered.

At a high level, defining policies in an EEM works as follows:

1. You define one or more policies.
2. The event detector software watches for the conditions that match those defined by the policy.
3. When a condition occurs, the event is passed to the event manager server.
4. The server invokes any policy that has registered for that particular event.
5. The actions that are defined within the policy are carried out.

There are several building blocks that make up an EEM applet, including events and action. EEM applets use a logic similar to the *if-then* statements used in many programming languages (for example, *if* an event happens, *then* an action is taken).

As shown in Example 17.1, a number of actions can be tied to an EEM applet.

EXAMPLE 17.1 **EEM Applet Actions**

```
R1(config-applet)# action test1 ?
  add                Add
  append             Append to a variable
  break              Break out of a conditional loop
  cli                Execute a CLI command
  cns-event          Send a CNS event
  comment            add comment
  context            Save or retrieve context information
  continue           Continue to next loop iteration
  counter            Modify a counter value
  decrement          Decrement a variable
  divide             Divide
  else               else conditional
  elseif             elseif conditional
  end                end conditional block
  exit               Exit from applet run
  file               file operations
  force-switchover   Force a software switchover
  foreach            foreach loop
  gets               get line of input from active tty
  handle-error       On error action
  help               Read/Set parser help buffer
  if                 if conditional
  increment          Increment a variable
  info               Obtain system specific information
  mail               Send an e-mail
  multiply           Multiply
  policy             Run a pre-registered policy
  publish-event      Publish an application specific event
  puts               print data to active tty
  regexp             regular expression match
  reload             Reload system
  set                Set a variable
  snmp-object-value  Specify value for the SNMP get request
  snmp-trap          Send an SNMP trap
  string             string commands
  subtract           Subtract
  syslog             Log a syslog message
  track              Read/Set a tracking object
  wait               Wait for a specified amount of time
  while              while loop
```

These are some of the most commonly used EEM applet actions:

- **action cli:** This action executes a Cisco IOS CLI command when an EEM applet is triggered.
- **action counter:** This action sets or modifies a named counter when an EEM applet is triggered.
- **action decrement:** This action decrements the value of a variable when an EEM applet is triggered.
- **action snmp-trap:** This action generates an SNMP trap when an EEM applet is triggered.
- **action mail:** This action sends a short email when an EEM applet is triggered.
- **action reload:** This action reloads a Cisco IOS device when an EEM applet is triggered.
- **action syslog:** This action writes a message to syslog when an EEM applet is triggered.
- **action put:** This action enables the printing of data directly to the local tty when an EEM applet is triggered.

ExamAlert

Before taking the ENCOR exam, be sure you understand the following recommendations for setting up an EEM applet.

Consider these pointers that assist in setting up an EEM applet:

- Include the **enable** and **configure terminal** commands at the beginning of the actions in an applet because the applet assumes that the user is in EXEC mode, not privileged EXEC or config mode.
- Use decimal labels similar to 1.0, 2.0, and so on when building applets. This makes it easier to insert new actions between other actions in the future as the need arises. For example, you could later insert 1.5 between 1.0 and 2.0 actions.
- Use the command **debug event manager action cli** to see the actions taking place when an applet is running.
- Use the command **debug event manager all** to show all the output for the configured actions while an applet is being executed.

Example 17.2 shows how to set up a simple syslog event detector. This example assumes that the device being used is an edge router, and GigabitEthernet 0/1 is Internet facing. When the interface GigabitEthernet 0/1 is shut down, the created applet can run to enable the interface. The applet is also set up to send an alert by email. The alert action at 3.5 is to send email with the body "R1 interface GigabitEthernet 0/1 is down."

EXAMPLE 17.2 **Creating an EEM Applet**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# event manager applet R1_interface_shutdown
R1(config-applet)# event syslog pattern "Interface
GigabitEthernetEthernet0/1, changed state to administratively down"
R1(config-applet)# action 1.0 cli command "enable"
R1(config-applet)# action 1.5 cli command "configure terminal"
R1(config-applet)# action 2.0 cli command "interface
GigabitEthernetEthernet0/1"
R1(config-applet)# action 2.5 cli command "no shutdown"
R1(config-applet)# action 3.0 cli command "end"
R1(config-applet)# action 3.5 mail server "10.10.10.1" to
"examcram@pearson.com" from "eem_example@pearson.com." subject
"interface_shutdown." body "R1 interface GigabitEthernet0/1 is down"
R1(config-applet)# end
R1#
```

This example shows the use of an EEM applet to call a Tcl script. Whereas Example 17.1 shows how to execute an action after a specific event occurs, this example shows how to manually execute an EEM applet, which then executes a Tcl script stored in the device's flash memory. When you call an EEM script, you manually trigger an applet, and there are no automatic events that the applet is monitoring. It only runs when you manually trigger it. In this case, the EEM script is configured with the command **event none** as it is not monitoring any events. To manually run the applet, you use the command **event manager run** *applet-name*. To view the Tcl script stored in flash memory, you use the command **more** *file-location*:*filename* (for example, **more flash:pingtest.tcl**).

For basic Tcl scripts, you can type line-by-line in the Tcl shell. To get to the Tcl interpreter, you can enter **tclsh** in privileged mode, as shown here, and then paste in the contents of the script:

```
R1#
R1# tclsh
R1(tcl)#
```

Example 17.3 shows the contents of a sample Tcl ping script that can be stored in flash memory.

EXAMPLE 17.3 **Contents of a Sample Tcl Script**

```
foreach address {
10.10.10.1
10.10.10.2
10.10.10.3
} { ping $address
}
```

Example 17.4 shows the running of a Tcl script.

EXAMPLE 17.4 **Running a Tcl Script**

```
R1#
R1# tclsh
R1(tcl)# foreach address {
+>(tcl)# 10.10.10.1
+>(tcl)# 10.10.10.2
+>(tcl)# 10.10.10.3
+>(tcl)# } { ping $address
+>(tcl)# }
```

You can also save this script in the device's flash memory. From there you can create an EEM applet to call the script. When the applet is executed, the router pings the IP addresses defined in the Tcl script. In Example 17.5, the **event none** command is used, which means the applet is not automatically monitoring an event; rather, the applet runs only when it is triggered manually.

EXAMPLE 17.5 **Creating an EEM Applet to Call Tcl Script**

```
event manager applet ExamCramPing
 event none
 action 1.0 cli command "enable"
 action 1.1 cli command "tclsh flash:/examcram_ping_script.tcl"
R1# event manager run ExamCramPing
```

Example 17.6 demonstrates the use of the action **syslog**. In this case, a syslog message is posted, indicating that the **no shutdown** command ran to bring GigabitEthernet 0/1 back up after the interface was shut down.

EXAMPLE 17.6 **Using the Syslog Action in an EEM Applet**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# event manager applet R1_interface_shutdown
R1(config-applet)# event syslog pattern "Interface
GigabitEthernetEthernet0/1, changed state to
administratively down"
R1(config-applet)# action 1.0 cli command "enable"
R1(config-applet)# action 1.5 cli command "configure terminal"
R1(config-applet)# action 2.0 cli command "interface
GigabitEthernetEthernet0/1"
R1(config-applet)# action 2.5 cli command "no shutdown"
R1(config-applet)# action 3.0 cli command "end"
R1(config-applet)# action 3.5 syslog msg "Due to Interface
GigabitEthernetEthernet0/1 shutdown, no shutdown command ran"
R1(config-applet)# end
R1#
```

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which EEM architecture component bridges the Cisco IOS subsystems used in the event detectors and policies?
   - ❍ **A.** EEM server
   - ❍ **B.** EEM detector
   - ❍ **C.** EEM applet
   - ❍ **D.** Tcl script

2. Which event detector publishes an event when the Cisco IOS **event manager run** CLI command manually executes an EEM policy?
   - ❍ **A.** CLI event detector
   - ❍ **B.** Syslog event detector
   - ❍ **C.** Tcl script
   - ❍ **D.** None event detector

# Answers

1. **A** is correct. The EEM server bridges the Cisco IOS subsystems used in the event detectors and policies.
2. **D** is correct. The none event detector publishes an event when the Cisco IOS **event manager run** CLI command manually executes an EEM policy.

# Review Questions

1. True or false: An EEM applet can assist in cases where you need a temporary workaround while you determine the root cause of an issue.
   - ❍ **A.** True
   - ❍ **B.** False

2. What command do you use to manually run an EEM applet?
   - ❍ **A. event manager run** *applet-name*
   - ❍ **B. event none** *applet-name*
   - ❍ **C. more** *file-location:filename*
   - ❍ **D. event manager action cli**

# Answers to Review Questions

1. **A** is correct. EEM scripts can assist in cases where you need a temporary workaround while you determine the root cause of an issue. For example, if you have a problem that is intermittent but that is fixed by an interface reset, you can use EEM scripts to trigger a reset as soon as the problem occurs so you can then manually intervene to find the root cause of the problem.
2. **A** is correct. To manually run an applet, you use the command **event manager run** *applet-name.*

# Further Reading

- ▶ *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*
- ▶ *CCNP and CCIE Enterprise Core & CCNP Advanced Routing Portable Command Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers configuration management and orchestration.

CHAPTER 18

# Configuration Management and Orchestration

**This chapter covers the following official ENCOR 350-401 exam objective:**

- 6.7 Compare agent vs. agentless orchestration tools, such as Chef, Puppet, Ansible, and SaltStack

Traditionally, network changes were made manually on individual Cisco network devices. For example, to add a handful of new VLANs or make firewall changes across the campus and data center, someone needed to log on to each device and make the changes manually. Today, we have configuration management tools that achieve the same results faster and more reliably. When you have a defined manual workflow to perform a set of tasks, it is beneficial to use configuration management tools to automate the tasks. Open-source tools such as Puppet, Chef, SaltStack, and Ansible can dramatically reduce the number of manual interactions with the network.

An automation framework—such as Puppet, Chef, SaltStack, or Ansible—incorporates idempotency, meaning ensuring that if no change is needed, the system ends in that same state. Basically, it makes no change if the system is already in that state. These configuration management tools can significantly help you automate applications, infrastructure, and networks to a high degree without doing any manual programming. In addition to reducing the time it takes to perform specific tasks, they offer greater predictability. You can reduce the risks related to manual misconfiguration by implementing proven and tested automation methods.

A number of common network-related tasks may benefit from the speed and consistency of configuration and orchestration tools. For example, you can use configuration management and orchestration tools to change device names/IP addresses, access control list (ACL) entries, usernames/passwords, SNMP settings, Syslog settings, and routing protocol configurations.

This chapter covers the operation and use of common configuration management and orchestration tools, such as Puppet, Chef, SaltStack, and Ansible. It examines these tools from a high-level perspective and provides some examples of the many options available with these tools. It is divided into two sections. The first section covers the operation and use of the agent-based

orchestration tools Puppet, Chef, and SaltStack. The second section looks at the operation and use of the agentless orchestration tools Ansible and Puppet Bolt. While you can use tools like Python to do manual programming on some Cisco platforms, the tools we look at in this chapter are the more commonly used network configuration management and orchestration tools. For the ENCOR 350-401 exam, you should at least understand how to use these tools to do basic manipulation of the configuration of Cisco network devices.

**This chapter covers the following technology topics:**

- Agent-Based Orchestration Tools
- Agentless Orchestration Tools

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. Which agent-based configuration management tool is based on Ruby and is similar in operation to Puppet?
2. Which agentless configuration management tool uses Playbooks?

### Answers

1. Chef
2. Ansible

# Agent-Based Orchestration Tools

This section covers several agent-based orchestration tools, including their main features and use cases. *Agent-based* means a software agent is installed on all the devices that need to be managed by the tool. It is not possible to always load an agent on all Cisco network platforms. Thus, agent-based tools have a higher barrier to entry than agentless tools when it comes to getting started with automation. This section provides a number of examples and command snippets that can help you decide which tools are best suited for your environment and needs.

## Puppet

Puppet, which was created in 2005, is one of the oldest agent-based orchestration tools. Cisco supports Puppet on a number of platforms, including Catalyst and Nexus switches and Unified Computing System (UCS). Puppet, which is written in Ruby, models the desired system states, enforces those states, and reports variances so you can track what it is doing. To model system states, Puppet uses a declarative resource-based language; this means a user describes the desired final state (for example, "this VLAN must be present" or "this route must be present") rather than describing a series of steps to execute.

You can use the declarative, readable Puppet domain-specific language (DSL) to define the desired end states of an environment. Puppet then converges the

infrastructure to the desired state. If you have a predefined configuration that every new switch should receive, the Puppet intent-based automation solution can quickly automate those repetitive configuration management tasks.

**ExamAlert**

For the ENCOR exam, make sure you understand the components and concepts of Puppet.

Puppet provides an intent-definition construct called a *manifest*. When you deploy a manifest on a Cisco network device, it translates into network configuration settings and commands for collecting information from the device.

Puppet includes the following components:

- **Puppet agent:** A Puppet agent runs on a managed device (node). The Puppet agent connects to the Puppet master periodically.
- **Puppet master:** The Puppet master typically runs on another dedicated server and serves multiple devices. The Puppet master compiles and sends a configuration manifest to the Puppet agent. The agent then reconciles the manifest with the node's current state and updates the state based on the differences.
- **Puppet database:** The Puppet database stores changes or automation tasks. It can be the same server as the Puppet master server or a different system.

Puppet has *modules* that are used for the configuration of almost anything that can be configured manually on a device. Puppet includes the following concepts:

- **Manifest:** This is the code that configures the clients or nodes running the Puppet agent. Manifests are pushed to devices using SSL and require certificates to be installed to ensure the security of the communications between the puppet master and the puppet agents. Manifest files have the extension .pp.
- **Module:** This is a collection of manifests, templates, and files that are used to configure the options of a specific feature. Modules can be reused. The module used in this chapter is called cisco_ios.

- **Resource:** This is the part of the Puppet code that describes a particular aspect of a system.
- **Class:** This is a collection of resources that you can reuse multiple times.
- **Catalog:** This is a compiled configuration for each individual node.

Automation of various network configuration and management tasks from a Puppet server enables you to dramatically reduce configuration time while eliminating manual tasks that are repetitive and error-prone. For example, the provisioning of network constructs like VLANs, ports, network routes, QoS parameters, and access control can be easily automated with Puppet. Further, Puppet can make many ongoing management operations dramatically easier, including configuration management, compliance auditing, and monitoring.

Example 18.1 shows a simple manifest file that configures the message-of-the-day (MOTD) banner on a Cisco Catalyst switch.

EXAMPLE 18.1 **MOTD Manifest File**

```
banner { 'default':
   motd => 'Unauthorised Access to this Device is Strictly Prohibited.
Violators will be prosecuted',
}
```

# Chef

Chef is another agent-based configuration management tool. Chef is built around a couple simple concepts: achieving a desired state and a centralized modeling of IT infrastructure. Chef enables you to quickly manage almost any infrastructure. Chef is based on Ruby, uses a declarative intent-based model, is agent-based, and refers to its automation instructions as *recipes*. A recipe defines a reusable set of configuration or management tasks. A group of recipes is referred to as a *cookbook*. Chef allows a recipe to be deployed on numerous network devices. When deployed on a Cisco network device, a recipe translates into a network configuration or a set of commands for gathering statistics and analytics information.

ExamAlert

For the ENCOR exam, make sure you understand the components and concepts of Chef.

Chef automation involves the following important components and concepts:

- **Chef server:** The Chef server functions as a hub for configuration data. It stores the following:
  - **Cookbook:** A group of recipes
  - **Recipes:** Policies that are applied to nodes
  - **Metadata:** Descriptions of each registered node that is being managed by a Chef client
- **Node:** A node is a network device that is being configured to be maintained by a Chef client.
- **Chef client:** A Chef client runs locally on the nodes that are registered with the Chef server. It carries out all configuration tasks specified by the run list and brings the client into the desired state.
- **Chef resources:** A Chef resource is a group of managed objects/attributes and one or more corresponding implementations. It defines the desired state for a configuration item and declares the steps needed to bring that item to the desired state. The core layers of a resource are as follows:
  - **Resource type:** A definition of a managed object
  - **Resource provider:** The layer that is responsible for the implementation of management tasks on objects
- **Cookbook:** In Chef, a cookbook defines a scenario and contains everything that is required to support that scenario. It is used for device configuration and policy distribution, and it includes the following:
  - Recipes that specify the resources to use and the order in which they are to be applied
  - Attribute values
  - File distributions
  - Templates
  - Extensions to Chef, such as libraries, definitions, and custom resources
- **Recipe:** In Chef, a recipe is a collection of resources defined using patterns (that is, resource names, attribute/value pairs, and actions). A recipe is subject to the following:
  - Must be stored in a cookbook
  - May use the results of a search query and read the contents of a data bag

- May have a dependency on one or more recipes
- Must be added to a run list before it can be used by the Chef client
- Is always executed in the same order as listed in a run list
- Permits the Chef client to run a recipe only when asked

Example 18.2 shows a simple cookbook and the configuration of two interfaces. The first interface is configured as a Layer 3 interface (due to the **switchport mode 'disabled'** option). The second interface is configured as an access port in VLAN 10 with the **switchport mode 'access'** and **access vlan 10** commands.

EXAMPLE 18.2 **Chef Cookbook with Interface Configurations**

```
cisco_interface 'GigabitEthernet0/1' do
  action :create
  ipv4_address '10.10.10.1'
  ipv4_netmask_length 24
  ipv4_proxy_arp true
  ipv4_redirects true
  shutdown true
  switchport_mode 'disabled'
end

cisco_interface 'GigabitEthernet0/2' do
  action :create
  access_vlan 10
  shutdown false
  switchport_mode 'access'
  switchport_vtp true
end
```

# SaltStack

SaltStack—commonly called just Salt—is an open-source Python-based automation software. It was originally developed by Thomas S. Hatch and released in 2011. Salt uses a modular design and uses Python modules to handle aspects of the available Salt systems. Salt is a publisher/subscriber model, which means the Salt master publishes jobs that need to be executed, and Salt minions subscribe to those jobs. If a job applies to a minion, the minion executes it and reports back data to the Salt master.

Through its highly modular and extensible design, Salt can be easily molded for a number of networking use cases. The functions within Salt support both

remote execution and configuration management. There are many types of Salt modules that manage specific actions. These modules can be added to systems that support dynamic modules. These modules, which help manage portability, make up the core API of system-level functions that Salt systems use.

Depending on the Cisco platform, Salt can run on a switch or off a switch or router platform. For example, in the Nexus platform, the Cisco NX-OS Guest Shell hosts SaltStack minions and provides automated orchestration of one or more switches through a unified interface. In this chapter, we are concerned with the off-switch option. With this option, the Salt primary runs the Salt software on a network device and communicates through SSH (the SSH proxy minion). The Cisco network device then interprets the commands, performs required configuration tasks, and reports success or failure to the appropriate proxy minion. The proxy minion, in turn, transmits this data back to the Salt primary. This off-switch implementation of Salt is called *Salt SSH*, and it can technically be considered agentless.

**ExamAlert**

For the ENCOR exam, make sure you understand the main components of Salt as well as their functions.

These are the main components of Salt:

- **Salt minion:** A Salt minion executes the instructions sent by the Salt master, reports on job success, and provides data related to the underlying host.
- **Proxy minion:** A proxy minion runs intermediate software between the Salt master and the Salt minion.
- **Salt master:** A Salt master is a master daemon for sending configurations and commands to the Salt minions. It is a machine that manages the Salt automation infrastructure and dictates policies. Salt minions, in turn, receive the configuration and commands from the master daemon.
- **Beacon:** Beacons are monitoring tools that can listen for a number of system processes on Salt minions. They can then trigger reactors to implement a change or troubleshoot an issue. Beacons are used for an array of purposes, including automated reporting, error log delivery, and resource monitoring.
- **Reactor:** Reactors expand Salt with automated responses using prewritten remediation states. When used with reactors, beacons can create

automated prewritten responses to infrastructure and application issues. Reactors can be used in a number of scenarios, including scaling infrastructure, notifying network administrators, restarting failed services, and performing automatic rollbacks.

- **Target:** A target is a group of minions that span one or many masters that a job's Salt command applies to. The Salt master indicates which minions should execute a particular job by defining a target.
- **Grain:** When you target a specific minion by its ID or target minions by their shared traits or characteristics, these are referred to as grains.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which of the following is not a component of Puppet?
   - ❍ **A.** Puppet master
   - ❍ **B.** Puppet agent
   - ❍ **C.** Puppet target
   - ❍ **D.** Puppet database

2. Which Chef component defines a scenario and contains everything that is required to support that scenario?
   - ❍ **A.** Recipe
   - ❍ **B.** Cookbook
   - ❍ **C.** Server
   - ❍ **D.** Client

## Answers

1. **C** is correct. The Puppet master, agent, and database are the three high-level components that make up the Puppet automation toolkit. A target is not a component.
2. **B** is correct. A Chef cookbook defines a scenario and contains everything that is required to support that scenario. A cookbook is used for device configuration and policy distribution.

# Agentless Orchestration Tools

This section covers two agentless orchestration tools that you can use for configuration management of network devices: Ansible and Bolt.

## Ansible

Ansible is an agentless software platform written in Python that was created as an alternative to Puppet and Chef. Ansible is used for deployment and configuration management of networking and compute infrastructure. It uses *playbooks*, which are basically scripts that run against devices in the environment that is being managed. A playbook features a state-driven resource model that describes the desired state of managed systems and services. You can use Ansible to automate the configuration of compute and switching resources in an agentless manner.

Ansible is a straightforward and powerful tool for intent-based network automation. Because it is agentless, Ansible lowers the barriers to entry to getting started with automation. Ansible can integrate and automate using any API. For example, integrations can use NETCONF, REST APIs, SSH, and SNMP, if needed.

Ansible has the following features, which are typical of orchestration software that uses an agentless architecture:

- A push-based model
- Scripts that run on the management server, connect to the managed device, and execute tasks
- No timer, as control lies with the management server

Ansible can be installed using the package manager on most popular Linux distributions, like Red Hat, CentOS, Fedora, Debian, and Ubuntu. Installation can be done via the Python package manager (pip) as well. To begin automating Cisco IOS XE devices, no additional effort is required as all Cisco IOS XE modules are included in Ansible Core. Once a username and a password are configured/provided, you can start managing devices using Ansible. You just have to ensure that the username provided with a playbook has the appropriate role privilege to allow device configuration changes. You also need SSH access to the Cisco IOS and IOS XE network devices so that a secure connection can be established.

**ExamAlert**

Before taking the ENCOR exam, make sure you know the steps that Ansible uses to connect to a network device to orchestrate configuration.

To get started with Ansible, you need just playbooks, a server configuration file, and an inventory file. Because Ansible is agentless, you do not need to install a software agent on the remote hosts to automate those devices; you just need SSH access. Ansible relies on SSH access to a device for executing commands. It also does not require the use of a dedicated server; you just need to install Ansible on a control station. You can have many machines with Ansible installed, and you can use them simultaneously to automate any given environment.

When you run a playbook, Ansible connects to devices according to the playbook and carries out the instructions from that file. It is important to understand the following Ansible concepts:

- **Play:** A play is a set of tasks for hosts or groups of hosts.
- **Task:** A task is an action that references a module to run with input arguments and actions.
- **Playbook:** A playbook specifies a list of tasks that runs in sequence across one or more managed network devices. Each task can also run multiple times, with a variable taking a different value. Playbooks use the Yet Another Markup Language (YAML) format, which is a simpler markup language than XML, and each includes one or more plays.
- **Inventory:** In Ansible, an inventory is the representation of information about a managed network device. It describes to which groups a host belongs and the properties those groups and hosts have. You can hierarchically organize groups.
- **Template:** A template enables you to generate configuration files from values that are set in various inventory properties. You can store one template in the source control, and it can apply to many different environments.
- **Role:** Roles give you a way to encapsulate common tasks and properties for reuse. Once you start writing the same tasks in multiple playbooks, you can turn them into roles for easy reusability.

- **Module:** A module is code that is used to perform an action on a managed device. A module can be written in a number of languages, such as Python.

Example 18.3 shows an Ansible host file with two groups and two Catalyst switches each to be managed within these groups.

EXAMPLE 18.3 **An Ansible Host File**

```
$ cat /etc/ansible/hosts

[ios-xe:vars]
ansible_network_os=ios
ansible_connection=network_cli

[catalyst3k]
catalyst3k1
catalyst3k2

[catalyst9k]
catalyst9k1
catalyst9k2

[ios-xe:children]
catalyst3k
catalyst9k
```

Example 18.4 shows a configuration to remove the MOTD banner and add a Login Banner.

EXAMPLE 18.4 **Configuring Ansible to remove the MOTD Banner and Add a Login Banner from a File**

```
- name: Configure login banner
  cisco.ios.ios_banner:
    banner: login
    text: |
      This is ExamCram Login Banner
    state: present

- name: Remove motd banner
  cisco.ios.ios_banner:
    banner: motd
    state: absent
```

```
- name: Configure banner from file
  cisco.ios.ios_banner:
    banner: motd
    text: "{{ lookup('file', './config_partial/raw_banner.cfg') }}"
    state: present

- name: Configure the login banner using delimiter
  cisco.ios.ios_banner:
    banner: login
    multiline_delimiter: x
    text: this is my login banner
    state: present
```

# Bolt

Bolt, which is part of Puppet, is an open-source orchestration tool that can help you automate your Cisco infrastructure. Bolt is agentless, so you can simply have it installed on a local workstation and connect directly to network devices over SSH or WinRM.

Bolt allows you to get up and running quickly by automating network devices. You can execute tasks against devices to perform ad hoc operations. For example, the cisco_ios module comes with tasks already defined to save the running configuration to the startup configuration or to execute a CLI command against network devices. Once you connect to an IOS XE device, you can run a task, apply a manifest, and run a plan against the device.

Bolt plans are sets of tasks that are combined with other logic. They give you the flexibility to do complex task operations, such as running multiple tasks with one command or running certain tasks based on the results of another task.

Example 18.5 shows how to use Bolt to quickly do some network diagnosis using ping. Network troubleshooting with Bolt is an ideal example of where you can use Bolt plans. A plan allows you to chain together multiple commands as part of a standard troubleshooting script.

EXAMPLE 18.5 **Troubleshooting with Bolt**

```
> bolt task run cisco_ios::cli_command --nodes cisco_ios command="ping
8.8.8.8" raw=true
Started on 1.1.1.1...
Finished on 1.1.1.1:
  ping 8.8.8.8
```

```
  Type escape sequence to abort.
  Sending 5, 100-byte ICMP Echos to 8.8.8.8, timeout is 2 seconds:
  .....
  Success rate is 0 percent (0/5)
  ios-switch#
  {
  }
Successful on 1 node: 1.1.1.1
```

# Configuration Management and Orchestration Tools Comparison

To wrap up this chapter, Table 18.1 shows a comparison of the automation tools covered in this chapter.

TABLE 18.1 **Comparison of Configuration Management and Automation Tools**

| | Puppet | Chef | SaltStack | Ansible |
|---|---|---|---|---|
| **Architecture** | Puppet master and puppet agents | Chef server and clients | Salt master and minions | Control station and remote hosts |
| **Language** | Ruby | Ruby | Python | Python |
| **Automation instructions** | Manifests | Recipes (Cookbooks) | Modules | Playbooks |
| **Agent-based/ agentless** | Agent-based (but has an agentless version called Bolt) | Agent-based | Agent (but has an agentless version called Salt SSH) | Agentless |

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which of the following agentless tools has a low barrier to entry for getting started with automation and uses a push model?

   ❍ **A.** Puppet

   ❍ **B.** Ansible

   ❍ **C.** Chef

   ❍ **D.** SaltStack

2. True or false: A playbook uses YAML format, a simpler markup language than XML, and includes one or more plays.

   ❍ **A.** True

   ❍ **B.** False

# Answers

1. **B** is correct. Ansible is agentless, has a lower barrier to entry, and uses a push model.
2. **A** is correct. A Playbook uses the Yet Another Markup Language (YAML) format, which is a simple markup language, and includes one or more plays.

# Review Questions

1. Which of the following automation tools are considered agentless? (Choose two.)
   - ❍ **A.** SaltStack
   - ❍ **B.** Ansible
   - ❍ **C.** Puppet Bolt
   - ❍ **D.** Chef
   - ❍ **E.** Puppet

2. Which of these configuration management tools are built on Ruby? (Choose two.)
   - ❍ **A.** Puppet
   - ❍ **B.** Chef
   - ❍ **C.** SaltStack
   - ❍ **D.** Ansible
   - ❍ **E.** Salt SSH

# Answers to Review Questions

1. **B** and **C** are correct. Ansible and Puppet Bolt do not need an agent installed on the managed network device and thus are considered agentless. Salt SSH (which is not an option in this question) is also considered agentless.
2. **A** and **B** are correct. Puppet and Chef are built on Ruby.

# Further Reading

- ▶ *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers enterprise network design principles.

CHAPTER 19

# Enterprise Network Design Principles

**This chapter covers the following official ENCOR 350-401 exam objectives:**

- 1.1 Explain the different design principles used in an enterprise network
- 1.1.a Enterprise network design such as Tier 2, Tier 3, and Fabric Capacity planning
- 1.1.b High availability techniques such as redundancy, FHRP, and SSO

This chapter looks at how the hierarchical network design model helps achieve high-performance, highly available, and scalable network designs. This chapter focuses on the different options for deploying an enterprise network architecture based on the hierarchical network design model. These options include the two-tier and three-tier models, and you will learn how they are used as part of an enterprise campus architecture for scaling from small LAN environments to large campus deployments. This chapter also examines how first-hop redundancy protocols (FHRPs), such as Hot Standby Router Protocol (HSRP), Virtual Router Redundancy Protocol (VRRP), and Gateway Load Balancing Protocol (GLBP), help you provide a redundant gateway for hosts. Finally, this chapter examines direct hardware redundancy mechanisms such as Stateful Switchover (SSO) and Nonstop Forwarding (NSF).

**This chapter covers the following technology topics:**

- Hierarchical LAN Design Model
- First Hop Redundancy Protocols (FHRPs)
- Hardware Redundancy Mechanisms

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. Which layer of the hierarchical LAN design model defines the summarization boundary for the network control plane protocols and serves as the policy boundary?
2. What is the primary motivation for deploying a two-tier (collapsed-core) design?
3. With HSRP, in the case of a tie, what is used to determine the priority?
4. Which hardware redundancy mechanism supports redundancy in a switch by allowing a redundant supervisor engine to take over if the primary supervisor engine fails?

### Answers

1. Distribution layer
2. The primary motivation for deploying a two-tier (collapsed-core) design is the cost reduction available along with most of the three-tier design benefits.
3. In the case of a tie, the primary IP addresses are compared, and the device with the higher IP address has priority.
4. Stateful Switchover (SSO)

# Hierarchical LAN Design Model

In a traditional flat network infrastructure deployment, all endpoints—including servers, workstations, and printers—connect using Layer 2 switches. All devices in a flat network share the same subnet and thus share the available bandwidth and are in the same broadcast domain. A broadcast packet sent from one device uses CPU time on every device in that broadcast domain. This setup is not efficient, and as the number of devices increases, network performance is severely affected. This limitation gave rise to the need for a hierarchical design model.

The hierarchical LAN design model divides a campus network into subsystems or building blocks. By assembling these building blocks or modules into a clear order, you can achieve a higher degree of stability, flexibility, and manageability for the individual pieces of the campus and the campus as a whole. A hierarchical design employs four fundamental design principles: hierarchy, modularity, resiliency, and flexibility. The modularity of the design can be easily replicated to other parts of the network, thus simplifying the network and providing an easy way to scale the network. This aids in providing a consistent deployment method to other parts on the network.

The hierarchical LAN design model's fundamental principles are designed so that each element in the hierarchy has a specific set of functions and services that it offers and a specific role in the design. The campus LAN environment was traditionally defined as a three-tier hierarchical model that consistsed of the core, distribution, and access layers. However, in small campus environments, the network may have only two tiers, with the core and distribution elements combined into one physical switch (distribution and core functions implemented in a single device).

Each layer in the hierarchical model focuses on specific functions and allows specific features and services to be deployed at a particular layer. This modular approach allows for flexibility in the network design and facilitates more straightforward implementation and troubleshooting. The access layer provides switches for the endpoint to connect to the network. The distribution layer acts as an aggregation point for the access layer and implements policies regarding security, traffic loading, and routing. The core layer provides connectivity between the distribution layer switches in larger networks. The hierarchical LAN design modular blocks are highlighted in more depth in the following sections.

## Access Layer

The access layer, also referred to as the network edge, is the first tier or edge of a campus network and is one of the most feature-rich parts of the network. It is where end devices (workstations, printers, cameras, and the like) attach to the wired portion of the network. It is also the layer where devices that extend out the network are connected (for example, IP phones and wireless access points [APs]).

The access layer provides the demarcation between the network infrastructure and the devices that connect and leverage that infrastructure. Thus, it provides security, quality of service (QoS), and a policy trust boundary. In the network security architecture, it is the first layer of defense and the first point of negotiation between end devices and the network infrastructure. The access layer therefore plays a crucial role in ensuring that a network is protected from malicious attacks.

The access layer can be segmented using technologies such as virtual LANs (VLANs) to allow different devices to go into different logical networks for performance, management, and security reasons. The access layer switches do not have interconnections to other access layer switches in the hierarchical LAN design model. Communication between endpoints on different access layer switches happens a level up, at the distribution layer.

## Distribution Layer

The distribution layer acts as a service and control boundary between the access layer and the core layer. It provides the boundary between the Layer 2 domain of the access layer and the Layer 3 domain of the core.

The distribution layer is the aggregation point for all access layer switches and is an integral part of the access-distribution block, providing connectivity and policy services for traffic flow within the access-distribution block. The distribution layer also provides the aggregation, policy control, and isolation demarcation point between the campus distribution building block and the rest of the network.

The distribution layer also defines the summarization boundary for the network control plane protocols (OSPF, EIGRP, Spanning Tree, and so on). It serves as the policy boundary between the devices and data flows with the access-distribution block and the rest of the network. In providing all of these functions, the distribution layer participates in both the access-distribution block and the core layer. The configuration choices for features in the distribution layers are often determined by the requirements of the access layer or the core layer.

## Core Layer

The core layer is the backbone that glues together all the campus architecture elements. The core layer also serves as the aggregator of all the other campus blocks and ties together the campus with the rest of the network. It is

responsible for providing scalability, high availability, and fast convergence to the network. Its main objective is to switch packets with minimal processing as quickly as possible.

The core layer also provides high-speed interconnectivity for end-user endpoints in the access layer and other network blocks. These network blocks may include the data center, the private cloud, the WAN, the Internet edge, or other areas within the network.

# Enterprise Network Architecture Options

When deploying a campus network, there are many enterprise network architecture design options available. The design that is used depends on the size of the campus and the resiliency, reliability, availability, performance, scalability, and security requirements. Because the campus network is modular, a mixture of design options can be used, depending on the business and technical requirements. In the following sections, we look at the designs most commonly used in the campus environment.

## Three-Tier Design

The three-tier design permits traffic aggregation and filtering at three successive routing and switching levels: the dedicated core, distribution, and access layers. (In contrast, the core layer is collapsed into the distribution layer in the collapsed-core design.) The three-tier design is often used to scale larger networks and is typically recommended when two or more pairs of distribution switches are required (especially in designs with multiple buildings where each building serves as a distribution layer). The three-tier design with a dedicated core layer allows a campus network to accommodate a higher growth rate without compromising the distribution block's design, the data center, or other parts of the network.

Without a core layer, all of the distribution layer switches need to be fully meshed. This design makes scaling the environment difficult due to the increased cabling requirement since each new building distribution switch needs full-meshed connectivity to all the distribution switches. Also, the routing complexity of having a full-meshed design increases as new neighbors are added.

Figure 19.1 shows the three-tier hierarchical model.

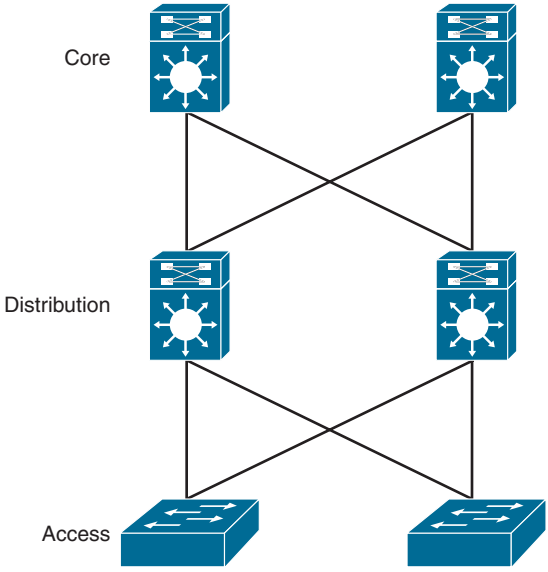



FIGURE 19.1 **Three-Tier Hierarchical Model**

## Two-Tier Design (Collapsed Core)

In a smaller campus, the network may have a two-tier switch design, where the core and distribution elements are combined into one physical switch infrastructure. The primary motivation for a two-tier design is to reduce network cost while maintaining most of the three-tier design benefits. The collapsed-core design must be able to provide the following services:

- High-speed physical and logical paths
- Layer 2 aggregation and demarcation point
- Routing and network access policies
- Intelligent network services (for example, QoS, network virtualization)

If a two-tier design is implemented, the collapsed-core switches must have enough capabilities and redundancy built in because they will be connecting multiple network blocks. The collapsed-core layer will be providing connectivity to the WAN edge block, the Internet edge block, the data center, and so on. On top of this, the collapsed-core layer switch still needs to provide LAN aggregation to the end-user access layer.

Figure 19.2 shows the two-tier hierarchical model (collapsed core).

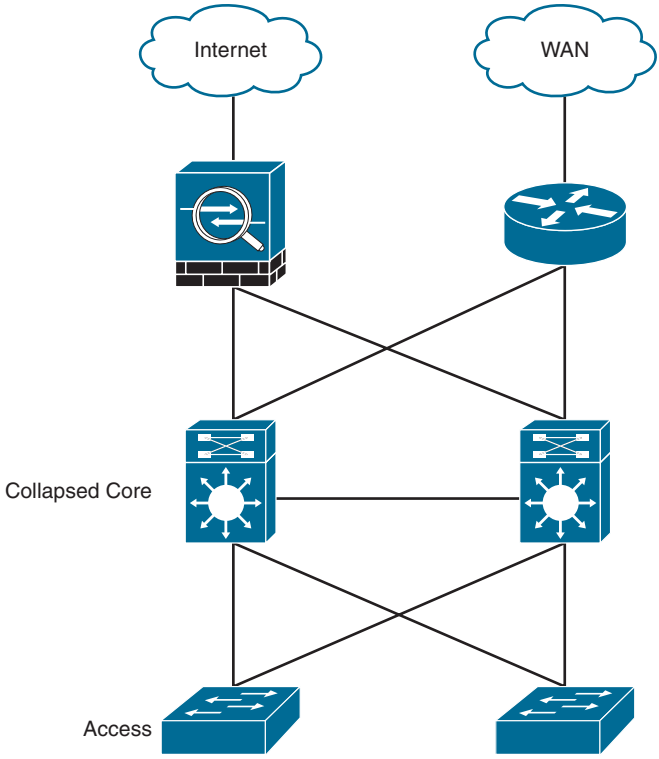



FIGURE 19.2 **Two-Tier Hierarchical Model (Collapsed Core)**

## Layer 2 Access Design

In a Layer 2 access design, all access switches are configured to run in Layer 2 forwarding mode, and the distribution switches are configured to run both Layer 2 and Layer 3 forwarding. VLAN-based trunks are used to extend the subnets from the distribution switches to the access layer. A default gateway or FHRP, such as VRRP, HSRP, or GLBP, provides redundant gateways to the access layer's end-user devices.

The Layer 2–only network design is the traditional LAN design and is normally cheaper than the other design options. However, one of the key disadvantages of a Layer 2 design is that the network is underutilized due to the way Layer 2 protocols are designed to build loop-free network topologies. With the access layer connecting directly to the distribution layer, Spanning Tree Protocol blocks low-priority ports, basically cutting the available bandwidth in half. EtherChannel helps you mitigate this inefficiency by aggregating the physical

ports into logical port channels. However, a newer design option, involving routing all the way down to the access layer, provides increased uplink utilization and faster convergence.

Figure 19.3 shows a Layer 2 access design.

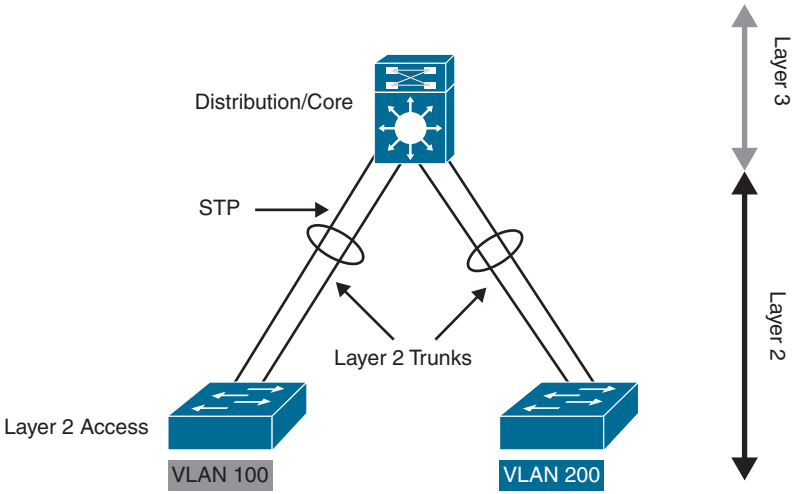



FIGURE 19.3 **Layer 2 Access Design**

## Layer 3 Access Layer (Routed Access) Design

The routed access layer design moves the boundary between Layer 2 and Layer 3 from the distribution layer to the access layer. Routing in the access layer simplifies configuration, optimizes distribution performance, and offers common end-to-end troubleshooting tools (such as **ping** and **traceroute**). Routing at the access layer removes the need for Layer 2 trunk links, which are replaced with point-to-point Layer 3 interfaces in the distribution layer.

At the network edge, Layer 3 access switches provide an IP gateway and become the Layer 2 demarcation point for locally connected endpoints that can be logically segmented by VLANs. The following benefits can be achieved by implementing a routed access design:

- It eliminates the need for Spanning Tree Protocol in the distribution layer. As a best practice, Spanning Tree Protocol enhancements, such as BPDU Guard and Port Fast, should still be implemented to harden and enhance the access layer.

- It shrinks the Layer 2 domain, thereby minimizing the number of endpoints affected by a denial-of-service (DoS)/distributed denial-of-service (DDoS) attack.
- It improves Layer 3 uplink bandwidth efficiency by suppressing Layer 2 broadcast at the access edge port.
- It eliminates the need for an FHRP, such as HSRP or GLBP.
- It enables easier troubleshooting, using standard end-to-end troubleshooting tools like **ping** and **traceroute**.
- It uses fast-converging routing protocols, such as OSPF and EIGRP.
- It increases uplink utilization between the distribution and access layers, thus increasing the bandwidth available to the end-user endpoints connected to the access layer switches.
- It improves performance by reducing resource utilization in the collapsed core/distribution layer. This enhancement is due to the fact that the collapsed core/distribution switch does not need to consume a lot of CPU cycles while processing a large number of MAC and ARP discoveries for end stations. Routed access reduces the load of Layer 2 processing and storage in the distribution layer by moving the load to Layer 3 access switches.

Figure 19.4 shows a Layer 3 access design.

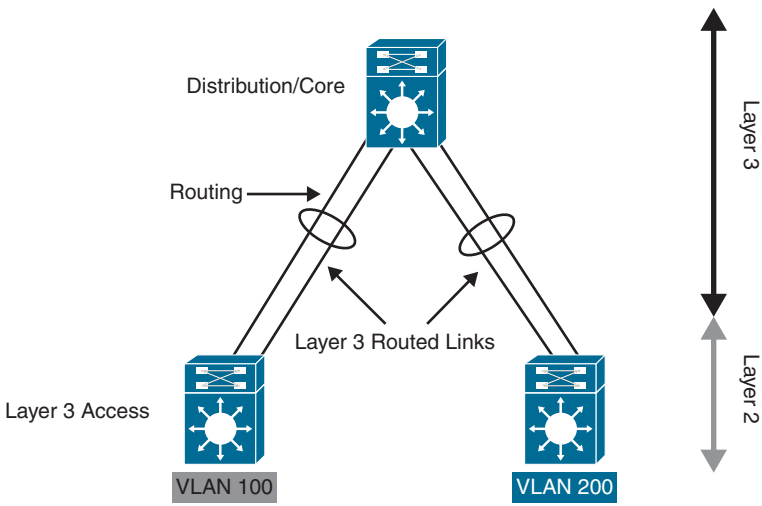



FIGURE 19.4 **Layer 3 (Routed Access) Design**

# Simplified Campus Design (Using VSS and StackWise)

The simplified campus design uses switch clustering technologies, such as Virtual Switching System (VSS) and stacking technologies, such as Cisco StackWise. The design utilizes these technologies to combine multiple physical switches to make them act as a single logical switch. VSS and StackWise can be applied at any of the campus building blocks to simplify them even further.

With the introduction of VSS and StackWise virtual switch concepts, a distribution switch pair can now be configured to run as a single logical switch. By converting the redundant physical distribution switches into a single logical switch, a significant change is made in the topology of the network. Rather than an access switch connecting with two uplinks to the distribution switches and needing a control protocol to determine which of the uplinks to use, the access switch connection can be simplified. The access switch can use a single Multichassis EtherChannel (MEC) upstream link to connect to a single distribution switch.

Figure 19.5 shows the VSS physical and logical topology.

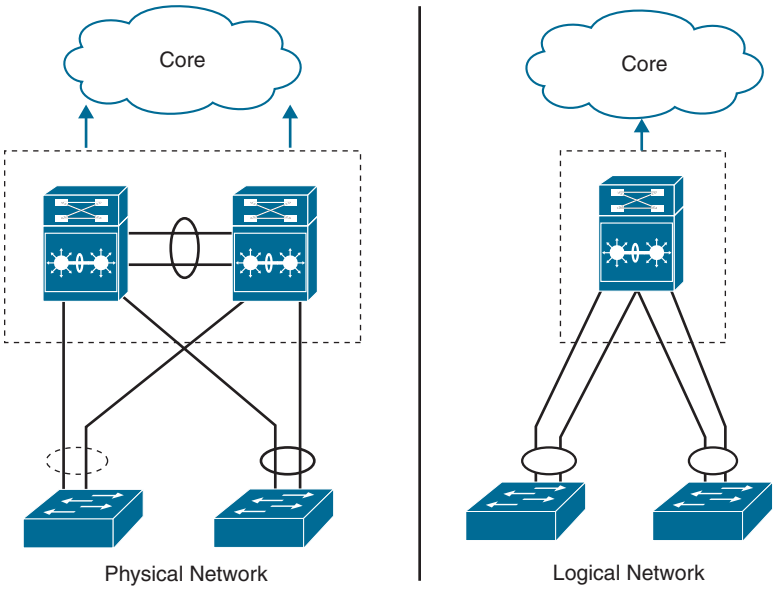



FIGURE 19.5 **VSS Physical and Logical Topology**

Implementing a simplified access design offers several advantages:

- It eliminates the need for an FHRP, such as HSRP or GLBP, because the IP gateway is on a single logical switch.
- There is a reduced dependency on Spanning Tree Protocol because EtherChannel automatically redistributes traffic to the remaining links in a bundle after a link failure. There is no need to wait on Spanning Tree Protocol to converge; in addition, Spanning Tree Protocol is already blocking 50% of all uplinks.
- Switch provisioning and maintenance are simplified because there are fewer physical switches to manage as they are combined into a logical design.
- Troubleshooting is simplified because the topology from the distribution layer to the access layer is a hub-and-spoke topology, which reduces the complexity of the design.

Figure 19.6 illustrates a traditional Spanning Tree Protocol topology and a VSS topology.

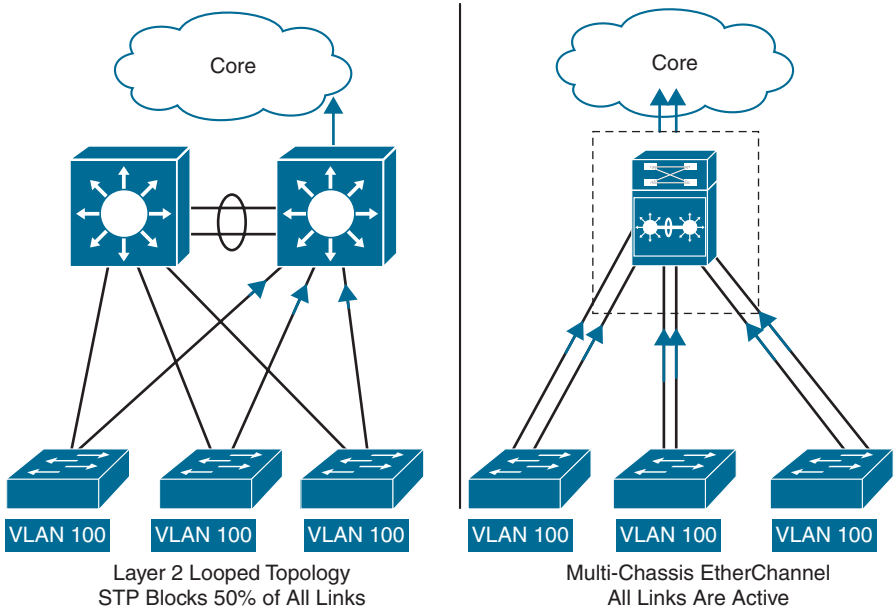



FIGURE 19.6 **STP and VSS Topologies**

## Software-Defined Access (SD-Access)

Cisco's Software-Defined Access (SD-Access) is the industry's first intent-based networking solution for enterprises. The solution provides policy-based automation from the edge to the cloud, with secure segmentation enabled through a single network fabric. By automating policy enforcement, SD-Access reduces the time it takes to adapt the network,improves issue resolution, and reduces the impact of security breaches.

SD-Access is part of the Cisco Digital Network Architecture (DNA) solution, which has caused an exponential and fundamental shift in how networks are designed, built, and managed. DNA allows enterprises to reduce operating expenditures (opex) and risks while creating infrastructures with consistent policies and services across and over wired, wireless, and hybrid networks.

SD-Access is managed using Cisco DNA Center, which is the controller for Cisco DNA-based networks. It provides a centralized software dashboard for managing enterprise networks. DNA Center uses intuitive workflows to simplify the provisioning of user access policies and provides advanced assurance capabilities. SD-Access is covered in more detail in Chapter 23, "SD-Access."

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which layer of the hierarchical LAN design model is the backbone that glues together all the campus architecture elements?
   - ❍ **A.** Core layer
   - ❍ **B.** Distribution layer
   - ❍ **C.** Access layer
   - ❍ **D.** Network edge

2. End-user devices should be connected to which layer of the hierarchical LAN design model?
   - ❍ **A.** Core layer
   - ❍ **B.** Distribution layer
   - ❍ **C.** Access layer
   - ❍ **D.** Network edge

3. Which layer of the hierarchical design model should provide the highest switching speed?
   - ❍ **A.** Core layer
   - ❍ **B.** Distribution layer
   - ❍ **C.** Access layer
   - ❍ **D.** Network edge

4. Virtual Switching System (VSS) and Cisco StackWise are part of which of the following enterprise network architecture designs?
   - ❍ **A.** Three-tier design
   - ❍ **B.** Simplified campus design
   - ❍ **C.** Software-Defined Access (SD-Access)
   - ❍ **D.** Two-tier design

# Answers

1. **A** is correct. The core layer is the backbone that glues together all of the campus architecture elements in the hierarchical LAN design model.
2. **C** is correct. The access layer is where end devices (workstations, printers, cameras, and so on) attach to the wired portion of the network.
3. **A** is correct. The core layer provides high-speed interconnectivity for end-user endpoints in the access layer and other network blocks.
4. **B** is correct. The simplified campus design uses switch clustering technologies, such as Virtual Switching System (VSS) and stacking technologies such as Cisco StackWise, to combine multiple physical switches to make them act as a single logical switch.

# First-Hop Redundancy Protocols (FHRPs)

In a traditional network design, to create a resilient IP gateway for the LAN segment, an FHRP must be used. An FHRP provides hosts in a segment with a consistent MAC and IP address for a default gateway that does not change when a failure occurs. Hot Standby Router Protocol (HSRP) and Virtual Router Redundancy Protocol (VRRP) are the most commonly used gateway redundancy protocols.

# Host Standby Router Protocol (HSRP)

HSRP is a Cisco-proprietary protocol that allows for the transparent failover of the first-hop device (default gateway). It is used in a group of routers or a multilayer switch interface/VLAN interface for selecting an active device interface/VLAN interface and a standby device interface/VLAN interface. In a group of device interfaces/VLAN interfaces, the active device is used for routing packets, and the standby device in the group takes over when the active device fails or when preset conditions are met.

HSRP allows a pair of routers or multilayer switches to work together in an HSRP group to provide a virtual IP address and an associated virtual MAC address. From the end-host perspective, the virtual IP address serves as the default gateway, and the end host learns the virtual MAC address via ARP. One of the routers in the group is the active router and is responsible for the virtual address. The other router in the group is in the standby state and monitors the active router. If the active router fails, the standby router assumes the active state. This is because the virtual addresses are always functional, regardless of which physical router is responsible for them. The physical interface/VLAN interface address cannot be used as an interface that the end hosts point to for their default gateway in HSRP. It needs to be a virtual IP address.

HSRP has two versions: HSRPv1 and HSRPv2. All devices in an HSRP group must have the same version number because HSRPv1 and HSRPv2 do not interoperate. (The hello messages are different.) Some of the differences between the HSRP versions are highlighted in Table 19.1.

TABLE 19.1 **HSRPv1 and HSRPv2 Comparison**

| HSRPv1 | HSRPv2 |
|---|---|
| Supports groups 0–255 | Supports groups 0–4095 |
| Supports IPv4 only | Supports IPv4/IPv6 |

| HSRPv1 | HSRPv2 |
|---|---|
| Multicast address is 224.0.0.2 | Multicast address is 224.0.0.102 |
| Virtual MAC address is 0000:0c07:acXX (with XX representing the HSRP group, in hexadecimal) | Virtual MAC address is 0000:0c9f:fXXX (with XXX representing the HSRP group, in hexadecimal) |
| This is the default version | Routers in a group need to be configured to run Version 2 with the **standby version 2** command. |

Devices configured in an HSRP group use three multicast messages:

- **Coup:** Coup messages are sent when a standby device wants to assume the active role.
- **Hello:** Hello messages are used to convey to other HSRP devices the device's HSRP priority and state information.
- **Resign:** Resign messages are sent by a device that is active when it is about to shut down or when a device that has a high priority sends a hello or coup message.

ExamAlert

For the ENCOR exam, make sure you know the different states of HSRP.

Devices configured for HSRP are always in one of the following states:

- **Active:** The device is responsible for forwarding (routing) packets that are being sent to it and responding to all ARP requests for the virtual IP address.
- **Init or disabled:** The device is not yet ready or able to participate in HSRP, possibly because the interface is not yet up. Locally configured groups with an interface that is down or groups without a specific interface IP address appear in the Init state.
- **Learn:** The device has not determined the virtual IP address and has not yet seen an authenticated hello message from the active device. In this state, the device still waits to hear from the active device.
- **Listen:** The device is receiving hello messages.
- **Speak:** The device is sending and receiving hello messages.
- **Standby:** The device is prepared to become the active device if the active device fails.

HSRP priority determines which device will be the active device and which device will be the standby device in the group. The device with the highest priority will be the active device, and the device with the second highest priority will be the standby device. If the priority is tied, the tie will be broken by the interface IP address. The device with the higher interface IP address will become the active device. The default priority is 100. HSRP preemption enables the HSRP router with the highest priority to immediately become the active router once it is available. When preemption is enabled, the switch with the highest priority will become the new active device. Preemption is disabled by default. Preemption can be enabled with the **standby preempt** command.

HSRP can be used to track objects or interfaces and decrement priority if the object or interface fails. Object tracking separates the tracking mechanism from HSRP and creates a separate tracking process that can be used by any other process as well as HSRP. The priority of a device can change dynamically when it is configured for object tracking and that object goes down. Some examples of objects that can be tracked are the line protocol state of an interface and the reachability of an IP route. If any of the objects that are being tracked goes down, the HSRP priority is reduced. The amount of decrease can be configured, but the default is 10. The purpose of object tracking is to increase availability and shorten the recovery time in the event that an object state goes down.

HSRP load sharing can be set up by configuring both paths to the core network with HSRP using Multigroup HSRP (MHSRP). For example, if you have two HSRP-enabled multilayer switches (MLSs) that participate in two separate VLANs, using 802.1q trunks, you use only one uplink toward the core of the network. This is especially true if you left the default HSRP priority values, hence causing the single MLS to become the active gateway for both VLANs.

To use both paths toward the core of the network, you can configure HSRP with MHSRP. For example, you can configure Group 10 for VLAN 10 and Group 20 for VLAN 20. For Group 10, one switch can be configured with a higher priority to become the active gateway, and the second switch can be configured to be the standby gateway. For Group 20, the second switch can be configured with a higher priority to become the active gateway, and the first switch becomes the standby gateway, achieving load sharing. The final configuration for a load-balanced HSRP setup like this would include configuring the spanning-tree root for the VLAN to be on the same switch as the active HSRP gateway for that VLAN.

HSRP authentication prevents rogue devices from joining the HSRP group. A rogue device may assume the active HSRP role and prevent hosts from communicating with the rest of the network, basically causing a DoS attack.

A rogue router may even forward all traffic and capture traffic from the hosts, basically achieving a man-in-the-middle attack. HSRP provides two types of authentication to mitigate these issues:

- **Plaintext authentication:** With plaintext authentication, a message matching the key configured on an HSRP peer is accepted. The maximum length of the key string is eight characters. The issue with plaintext messages is that they can be easily intercepted, so MD5-type authentication should be used, if available.
- **MD5 authentication:** With MD5 authentication, a hash is computed on a portion of each HSRP message and sent along with the HSRP message. When a peer receives a message and a hash, it performs hashing on the received message. If the received hash and the newly computed hash match, the message is accepted. MD5 should always be used, when available, because it is difficult to reverse the hash value itself, and the hash keys are never exchanged, even though the hash itself can be seen.

An HSRP hello message contains the priority of the router, the hello time, and hold time parameter values:

- **Hello time:** The hello time value indicates the interval between the hello messages that the router sends.
- **Hold time:** The hold-time value indicates how long the current hello time is considered valid.

The standby time includes an **msec** parameter to allow for subsecond failover within a group. The hello time and hold time are specified in seconds unless the **msec** keyword is used. By default, the hello time is 3 seconds (which means failover time could occur as much as 10 seconds before devices start to communicate with the new default gateway); the hello time can be an integer from 1 to 255. By default, the hold time is 10 seconds and can be an integer between 1 and 255. Lowering the hello time value results in increased traffic for hello messages, so it should be done cautiously.

The hello and hold timers need to match on both devices in a group and ideally should be set as low as possible to achieve fast convergence. Within milliseconds of the active router failing, the standby router can detect the failure, expire the hold time, and take over the active role. However, modification of the timer values must consider other parameters related to network convergence, such as dynamic routing protocol convergence time. Because the routing protocol may not have any awareness of HSRP configuration, suboptimal and black hole routing can occur if HSRP converges before the routing protocol.

# Virtual Router Redundancy Protocol (VRRP)

VRRP is a standard-based FHRP that serves as a standard-based alternative to the Cisco-proprietary HSRP. It is similar to HSRP in terms of its operation and configuration. The VRRP master role is analogous to the HSRP active role, and the VRRP backup role is analogous to the HSRP standby role.

A VRRP group has one master device and one or more backup devices. The device with the highest priority becomes the master, and the priority can range from 0 to 255. However, priority 0 has a special meaning: It indicates that the current master has stopped participating in VRRP. Setting the priority value to 0 helps trigger backup devices to quickly transition to the master role without waiting on the current master to time out.

Unlike HSRP, VRRP allows you to use an IP address of one of the physical VRRP group members as the virtual IP address. The master is the only device that sends advertisements (similar to HSRP hellos). The advertisements are sent to the multicast address 224.0.0.18 with protocol number 112. The default advertisement interval is 1 second, with a 3-second hold time; in comparison, HSRP has a default 3-second hello time and 10-second hold time.

Table 19.2 highlights the key differences between VRRP and HSRP.

TABLE 19.2 **VRRP and HSRP Comparison**

| VRRP | HSRP |
|---|---|
| Industry-standard based | Cisco proprietary |
| 1 master and several backups | 1 active, 1 standby, and several candidates |
| Virtual IP address can be the same as the physical IP address | Virtual IP address is different from the physical IP address |
| Uses 224.0.0.18 | Uses 224.0.0.2 and 224.0.0.102 |
| Can track only objects | Can track interfaces and objects |
| Default timer is 1-second hello time and 3-second hold time | Default timer is 3-second hello time and 10-second hold time |
| Authentication is no longer supported in RFC but still supported in Cisco IOS | Authentication supported |
| VRRP uses address 0000.5e00.01xx, where xx represents the group number. | HSRPv1 uses virtual MAC address 0000.0c07.acxx, where xx is the group number, and HSRPv2 uses 0000.0c9ffxxx, where xxx is the group number. |

In a campus environment, both HSRP and VRRP by default allow access switches to forward data out one of the uplinks to the distribution layer and require additional configuration for each distribution layer switch to allow

VLAN distributions across uplinks. Different HSRP or VRRP groups can be set up to provide redundancy; you can set up multiple groups and have different VLANs' traffic use different uplinks.

# Gateway Load Balancing Protocol (GLBP)

The third FHRP, Gateway Load Balancing Protocol (GLBP), provides greater uplink utilization for traffic exiting the access layer toward the distribution layer. This is achieved by balancing the load from the hosts across multiple uplinks.

GLBP is a Cisco-proprietary protocol, similar to HSRP, that protects data traffic from a failed device or circuit. However, GLBP provides true load balancing within a subnet/VLAN between a grouping of redundant devices. Similar to HSRP and VRRP, GLBP provides automatic device backup for IP hosts configured with a single default gateway on a LAN segment. With GLBP, multiple first-hop devices on the LAN combine to offer a single virtual first-hop IP device while sharing the IP packet forwarding load. Other devices on the LAN segment act as redundant GLBP devices that become active if any existing forwarding devices fail.

Although HSRP and VRRP function similarly to GLBP, with HSRP and VRRP, one member is elected as the active device to forward packets for the group's virtual IP address. The other devices in the group are redundant until the active device fails. In this case, the standby device has unused bandwidth that the protocol is not using. GLBP provides an advantage in that it allows for load balancing over multiple devices (gateways) using a single virtual IP address and multiple virtual MAC addresses. This translates to the forwarding load being shared among devices in the GLBP group rather than being handled by a single device while the others remain idle.

All hosts are configured with the same virtual IP address, and all the devices in the virtual device group participate in forwarding packets. Communication among GLBP members happens through hello messages sent every 3 seconds to multicast address 224.0.0.102, with UDP port 3222.

Table 19.3 highlights the key differences and similarities between GLBP and HSRP.

TABLE 19.3 **GLBP and HSRP Comparison**

| GLBP | HSRP |
|---|---|
| Cisco proprietary | Cisco proprietary |
| Active virtual gateway (AVG): 1 active, 1 standby, and several candidates. | 1 active, 1 standby, and several candidates |
| Active virtual forwarder (AVF): Multiple active and several candidates | |

| GLBP | HSRP |
|---|---|
| Virtual IP address is different from IP addresses on interfaces | Virtual IP address is different from real addresses |
| Can track only objects | Can track interfaces or objects |
| Uses 224.0.0.102 with UDP port 3222 | Uses 224.0.0.2 and 224.0.0.102 with UDP port 1985 |
| Authentication supported | Authentication supported |
| Default timers are 3-second hello time and 10-second hold time | Default timers are 3-second hello time and 10-second hold time |
| GLBP uses 0007.b400xxyy, where xx is the group number and yy is the AVF number. | HSRPv1 uses virtual MAC address 0000.0c07.acxx, where xx is the group number, and HSRPv2 uses 0000.0c9ffxxx, where xxx is the group number. |

Members within the GLBP group elect one gateway to be the active virtual gateway (AVG) for the group. Other members in the group provide backup for the AVG if the AVG becomes unavailable. The AVG assigns a virtual MAC address to each member of the GLBP group. Each gateway in the group assumes responsibility for forwarding packets sent to the virtual MAC address assigned to it by the AVG.

GLBP uses three different packet types. The hello packet is used for advertising protocol information. These packets are multicast and are sent when any virtual gateway or virtual forwarder is in the Speak, Standby, or Active state. The request and reply packets are used for virtual MAC assignment. They are both unicast messages to and from the AVG. These gateways are known as active virtual forwarders (AVFs). Finally, the AVG is also responsible for answering ARP requests for the virtual IP address. Load balancing is achieved by the AVG replying to the ARP requests with different virtual MAC addresses.

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. What types of authentication does HSRP support? (Choose two.)
   - ❍ **A.** Plaintext
   - ❍ **B.** SHA 256
   - ❍ **C.** MD5
   - ❍ **D.** SHA1

2. What are the possible states for a router or an MLS in an HSRP implementation? (Choose two.)
   - ❍ **A.** Master
   - ❍ **B.** Backup
   - ❍ **C.** Active
   - ❍ **D.** Standby

3. Which of the following FHRPs is a Cisco-proprietary implementation that allows multiple active forwarders to load balance traffic?
   - ❍ **A.** VRRP
   - ❍ **B.** GLBP
   - ❍ **C.** HSRP
   - ❍ **D.** MSTP

4. What multicast address does VRRP send its advertisement to?
   - ❍ **A.** 224.0.0.2
   - ❍ **B.** 224.0.0.102
   - ❍ **C.** 224.0.0.8
   - ❍ **D.** 224.0.0.18

# Answers

1. **A** and **C** are correct. HSRP supports plaintext and MD5 for authentication; MD5 authentication should always be used, when possible, since it hashes the password.
2. **C** and **D** are correct. Active and Standby are the two states for devices in an HSRP group.
3. **B** is correct. GLBP is a Cisco-proprietary protocol that protects data traffic from a failed device or circuit and allows load balancing between a group of redundant devices.
4. **D** is correct. VRRP sends advertisements to the multicast address 224.0.0.18 with protocol number 112.

# Hardware Redundancy Mechanisms

Cisco uses two mechanisms that help reduce the impact of specific network outages: Cisco Stateful Switchover (SSO) and Nonstop Forwarding (NSF), which build on the earlier Route Processor Redundancy (RPR) and Route Processor Redundancy Plus (RPR+). With redundant intra-chassis hardware (that is, redundant route processors) and the separation of the control and data plane, SSO and NSF make continuous packet forwarding with zero packet loss possible. When a hardware or software problem causes a route processor failure, links and interfaces remain up during switchover.

Chapter 27, "VRF Instances, GRE, and IPsec," covers control plane and data plane operations in depth. Here we provide just enough information on these topics to help you better understand hardware redundancy mechanisms that is covered in this section:

- **Control plane:** The control plane is where the Cisco switch or router learns about its environment, using various protocols to communicate with neighboring devices. The control plane is responsible for maintaining sessions and exchanging protocol information with other network devices. It includes dynamic routing protocols, ICMP, ARP, BFD, LACP, and so on.
- **Data plane:** The data plane is the forwarding plane and is responsible for the switching of packets through a network device. There could be features that affect forwarding in the data plane, such as QoS and ACLs. Typically, such traffic is processed by the device's hardware ASICs.

NSF relies on true separation of the data plane from the control plane during a supervisor switchover. The data plane continues to forward packets based on pre-switchover Cisco Express Forwarding (CEF) information. On the other hand, the control plane implements graceful restart routing protocol extensions to signal a supervisor restart to NSF-aware neighbors, reforms its neighbor adjacencies, and rebuilds the routing protocol database following a switchover. None of this would be possible without control plane and data plane separation.

# Stateful Switchover (SSO)

SSO supports redundancy in a switch by allowing a redundant supervisor engine to take over if the primary supervisor engine fails. SSO is frequently used with NSF to minimize the amount of time the network is unavailable to users following a switchover. (The switch continues to forward IP packets.)

SSO provides several useful benefits. Because the SSO feature maintains stateful information, user session information is maintained during a switchover. Line cards continue to forward network traffic with no loss to user sessions, thus improving network availability. In terms of enhancements over RPR, SSO provides faster switchover by fully initializing and configuring the standby route processor (RP); by synchronizing state information, routing protocol convergence time can be reduced. Finally, network stability can be improved due to the reduction in the number of route flaps created when routers in the network fail and thus take down their routing tables.

SSO establishes one of the RPs as the active processor, and the other is designated as the standby processor. SSO fully initializes the standby processor, and critical state information is synchronized between these two. During a switchover, the line cards do not reset; this provides a faster switchover between the processors. Several events can trigger a switchover:

- A hardware failure on the active supervisor engine
- A manual switchover or shutdown
- A clock synchronization failure between the supervisor engines

The critical point that should be noted is that Layer 2 traffic is not interrupted during a switchover. An SSO switchover preserves the forwarding information base (FIB) and adjacency entries and can forward Layer 3 traffic after a switchover.

You configure SSO by using the **redundancy** command and then, in redundancy mode, using the **sso** option.

Example 19.1 shows an SSO configuration on a Cisco Catalyst 4500X.

EXAMPLE 19.1 **Configuring SSO**

```
SW1
SW1(config)# redundancy
SW1(config-red)# mode ?
  rpr  Route Processor Redundancy
  sso  Stateful Switchover

SW1(config-red)# mode sso
SW1(config-red)# end
SW1#
```

You use the **show redundancy** command to verify that SSO is configured on the device, as shown in Example 19.2. Notice that the output shows the

configured and operating redundancy mode, current state, and uptime, among other information.

EXAMPLE 19.2 **Verifying SSO**

```
SW1#show redundancy
Redundant System Information :

------------------------------
       Available system uptime = 50 weeks, 4 days, 13 hours, 9 minutes
Switchovers system experienced = 0
              Standby failures = 0
        Last switchover reason = none

                 Hardware Mode = Duplex
    Configured Redundancy Mode = Stateful Switchover
     Operating Redundancy Mode = Stateful Switchover
              Maintenance Mode = Disabled
                Communications = Up

Current Processor Information :
------------------------------
               Active Location = slot 1/1
        Current Software state = ACTIVE
       Uptime in current state = 50 weeks, 4 days, 20 hours,
31 minutes
                 Image Version = Cisco IOS Software, IOS-XE Software,
Catalyst 4500 L3 Switch  Software (cat4500e-UNIVERSALK9-M), Version
03.09.02.E RELEASE SOFTWARE (fc4)
Technical Support: http://www.cisco.com/techsupport
Copyright (c) 1986-2017 by Cisco Systems, Inc.
Compiled Mon 01-May-17 02:17 by prod
        Configuration register = 0x2101

Peer Processor Information :
------------------------------
              Standby Location = slot 2/1
        Current Software state = STANDBY HOT
       Uptime in current state = 50 weeks, 4 days, 20 hours,
29 minutes
                 Image Version = Cisco IOS Software, IOS-XE Software,
Catalyst 4500 L3 Switch  Software (cat4500e-UNIVERSALK9-M), Version
03.09.02.E RELEASE SOFTWARE (fc4)
Technical Support: http://www.cisco.com/techsupport
Copyright (c) 1986-2017 by Cisco Systems, Inc.
Compiled Mon 01-May-17 02:17 by pr
        Configuration register = 0x2101

SW1#
```

You can use the **show redundancy clients** command to display a list of features that have been registered as SSO features (see Example 19.3).

EXAMPLE 19.3 **Using the show redundancy clients Command**

```
SW1#
SW1#show redundancy clients
Group ID =     1
     clientID = 20002    clientSeq =     4    EICORE HA Client
     clientID = 20001    clientSeq =    29    License Core HA Client
     clientID = 20011    clientSeq =    30    License Agent HA Client
     clientID = 20010    clientSeq =    31    SNMP SA HA Client
     clientID = 20007    clientSeq =    34    Installer HA Client
     clientID = 24701    clientSeq =    45    VMAN HA Client
     clientID =    29    clientSeq =    60    Redundancy Mode RF
     clientID =   139    clientSeq =    61    IfIndex
     clientID =  3300    clientSeq =    65    Persistent Variable
     clientID =    25    clientSeq =    71    CHKPT RF
     clientID =    77    clientSeq =    87    Event Manager
     clientID =   305    clientSeq =   110    Multicast ISSU
                                              Consolidation RF
     clientID =   304    clientSeq =   111    IP multicast RF Client
     clientID =    22    clientSeq =   112    Network RF Client
     clientID =    88    clientSeq =   113    HSRP
     clientID =    75    clientSeq =   129    Tableid HA
     clientID =   501    clientSeq =   136    LAN-Switch VTP VLAN
     clientID =    71    clientSeq =   138    XDR RRP RF Client
     clientID =    24    clientSeq =   139    CEF RRP RF Client
     clientID =  2900    clientSeq =   140    FFM RF
     clientID =   146    clientSeq =   141    BFD RF Client
     clientID =   301    clientSeq =   145    MRIB RP RF Client
     clientID =   306    clientSeq =   149    MFIB RRP RF Client
     clientID =   402    clientSeq =   160    TPM RF client
     clientID =   520    clientSeq =   161    RFS RF
     clientID =   210    clientSeq =   162    Auth Mgr
     clientID =     5    clientSeq =   163    Config Sync RF client
     clientID =   527    clientSeq =   170    Switch Stats Sync
     clientID =   502    clientSeq =   183    LAN-Switch Port Manager
     clientID =   200    clientSeq =   209    ETHERNET OAM RF
     clientID =   207    clientSeq =   211    ECFM RF
     clientID =   208    clientSeq =   214    LLDP
     clientID =    20    clientSeq =   226    IPROUTING NSF RF client
     clientID =   100    clientSeq =   229    DHCPC
     clientID =   101    clientSeq =   230    DHCPD
     clientID =    55    clientSeq =   237    GALIOS_CONFIG_SYNC
     clientID =    34    clientSeq =   242    SNMP RF Client
     clientID =    69    clientSeq =   248    AAA
     clientID =    35    clientSeq =   252    History RF Client
     clientID =   250    clientSeq =   276    EEM Server RF CLIENT
```

```
clientID =   252    clientSeq =   278    EEM POLICY-DIR RF CLIENT
clientID =    54    clientSeq =   280    SNMP HA RF Client
clientID =    57    clientSeq =   282    ARP
clientID =    50    clientSeq =   289    FH_RF_Event_
                                         Detector_stub
clientID =  2001    clientSeq =   290    CAT4K CHASSIS
clientID =  2002    clientSeq =   291    Link State
clientID =  2007    clientSeq =   292    Epm Switch
clientID =  2008    clientSeq =   294    K5Man
clientID =  2005    clientSeq =   297    Rkios
clientID =  2004    clientSeq =   298    GaliosModule
clientID =  1112    clientSeq =   304    Smart Install RF
clientID =   503    clientSeq =   306    Spanning-Tree Protocol
clientID =   507    clientSeq =   338    Switch Backup
                                         Interface client
clientID =   212    clientSeq =   340    REP Protocol
clientID =   105    clientSeq =   341    DHCP Snooping
clientID =  1510    clientSeq =   348    Call-Home RF
clientID =   151    clientSeq =   355    IP Tunnel RF
clientID =    94    clientSeq =   356    Config Verify RF client
clientID =   211    clientSeq =   357    Logging Redirect
clientID =   505    clientSeq =   360    Inline Power rf client
clientID =   516    clientSeq =   361    EnergyWise rf client
clientID =   506    clientSeq =   362    Igmp Snooping
clientID =   508    clientSeq =   363    Port Security
clientID =   509    clientSeq =   364    Switch IP Device
                                         Tracking
clientID =   515    clientSeq =   365    SISF table
clientID =   130    clientSeq =   371    CRYPTO RSA
clientID =   131    clientSeq =   372    PKI RF Client
clientID =   400    clientSeq =   374    IP Admission RF Client
clientID =  4005    clientSeq =   387    ISSU Test Client
clientID =    93    clientSeq =   391    Network RF 2 Client
clientID =   510    clientSeq =   394    LAN-Switch PAgP/LACP
clientID =  2006    clientSeq =   400    EbmHostMan
clientID =  2011    clientSeq =   401    EbmSpanMan
clientID =   141    clientSeq =   402    DATA DESCRIPTOR
                                         RF CLIENT
clientID =  1000    clientSeq =   412    CTS HA
clientID =  1001    clientSeq =   413    Keystore
clientID =  4022    clientSeq =   431    IOS Config SHELL
clientID =  4020    clientSeq =   432    IOS Config ARCHIVE
clientID =  4021    clientSeq =   433    IOS Config ROLLBACK
clientID =  4031    clientSeq =   434    ANCP
clientID =  4060    clientSeq =   436    ONEP
clientID =  4058    clientSeq =   438    PASSWD RF Client
clientID =   521    clientSeq =   473    MLD Snooping
clientID =   253    clientSeq =   476    DSENSOR
clientID =   254    clientSeq =   477    MSP
clientID =   255    clientSeq =   478    Flow Metadata
```

```
clientID =  2012    clientSeq =   479    DNS-AS Client
clientID =   526    clientSeq =   480    AutoQoS
clientID =   528    clientSeq =   492    MVR
clientID =   529    clientSeq =   493    VSL Sync
clientID = 24115    clientSeq =   495    mDNS
clientID = 11000    clientSeq =   498    EPC
clientID =  2010    clientSeq =   499    WCCPMan
clientID =  2013    clientSeq =   503    K5L3LispMan
```

# Nonstop Forwarding (NSF)

NSF can be used with SSO to minimize the amount of time a network is unavailable for users following a switchover. Enabling NSF high availability capabilities along with SSO informs the routers to continue maintaining the CEF entries for a short period and to continue forwarding packets during an RP failure until the control plane recovers.

NSF focuses on quickly rebuilding the routing information base (RIB) table after a supervisor switchover. The RIB is used for generating the FIB table for CEF, and then the FIB is downloaded to switch modules that can perform CEF. During a switchover, rather than waiting for Layer 3 routing protocols to reconverge, a router can use NSF to get assistance from other NSF-aware neighbors. The neighbors can then quickly provide routing information to the standby supervisor, allowing routing tables to be assembled quickly.

NSF must be built into the routing protocol on both the router that will need the assistance and the router that will provide the assistance. NSF is supported by CEF for forwarding and by BGP, EIGRP, IPv6, IS-IS, and OSPF for routing.

Some of the key prerequisites for NSF are as follows:

- The network devices to be configured for NSF must first be configured for SSO.
- For BGP NSF, all neighboring devices must be NSF aware and configured for graceful restart.
- For EIGRP NSF, all devices must be NSF capable or NSF aware; the NSF devices must be completely converged with the network before it can assist an NSF-capable device in an NSF restart operation.
- For IS-IS NSF, all neighboring devices must be NSF aware.
- For OSPF NSF, all networking devices on the same network segment must be NSF aware.
- For IPv6 NSF, IPv6 must be enabled on the network device.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. In which plane of operation does the learning of routing information for packet forwarding occur?
   - ❍ **A.** Management plane
   - ❍ **B.** Control plane
   - ❍ **C.** Service plane
   - ❍ **D.** Data plane

2. Which of the following features provides the fastest failover or supervisor or route processor redundancy?
   - ❍ **A.** RPR
   - ❍ **B.** RPR+
   - ❍ **C.** NSF
   - ❍ **D.** SSO

3. Which of the following features reduces the amount of time needed to rebuild the routing information after a switch supervisor module failure?
   - ❍ **A.** NSF
   - ❍ **B.** SSO
   - ❍ **C.** RPR
   - ❍ **D.** RPR+

## Answer

1. **B** is correct. A multilayer switch (MLS) or router learns from dynamic routing protocols how to make forwarding decisions in the control plane.
2. **D** is correct. SSO provides faster switchover by fully initializing and configuring the standby RP; by synchronizing state information, routing protocol convergence time can be reduced.
3. **A** is correct. NSF focuses on quickly rebuilding the RIB table after a supervisor switchover.

# Review Questions

1. Which of the following enterprise network architecture designs is also known as collapsed core?
   - ❍ **A.** Simplified campus design
   - ❍ **B.** Three-tier design
   - ❍ **C.** Layer 2 access layer
   - ❍ **D.** Two-tier design

2. Which of the following enterprise network architecture designs is typically used to scale larger networks and is typically recommended when two or more pairs of distribution switches are required?
   - ❍ **A.** Simplified campus design
   - ❍ **B.** Three-tier design
   - ❍ **C.** Layer 2 access layer
   - ❍ **D.** Two-tier design

3. Which of the following is the default hello time for HSRP?
   - ❍ **A.** 1 second
   - ❍ **B.** 3 seconds
   - ❍ **C.** 10 seconds
   - ❍ **D.** 9 seconds

4. Which of the following is the default hold time for VRRP?
   - ❍ **A.** 1 second
   - ❍ **B.** 10 seconds
   - ❍ **C.** 3 seconds
   - ❍ **D.** 30 seconds

5. Which plane of operation is the forwarding plane that is responsible for the switching of packets through a network device?
   - ❍ **A.** Management plane
   - ❍ **B.** Control plane
   - ❍ **C.** Service plane
   - ❍ **D.** Data plane

# Answers to Review Questions

1. **D** is correct. The two-tier enterprise network architecture design is also known as the collapsed core.
2. **B** is correct. The three-tier enterprise network architecture design is typically used to scale larger networks and is typically recommended when two or more pairs of distribution switches are required.
3. **B** is correct. By default, the hello time for HSRP is 3 seconds.
4. **C** is correct. By default, the hold time for VRRP is 3 seconds (that is, 3 times the 1-second hello time).
5. **D** is correct. The data plane is the forwarding plane that is responsible for the switching of packets through a network device.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers WLAN deployments.

CHAPTER 20

# Wireless LAN Deployments

**This chapter covers the following official ENCOR 350-401 exam objectives:**

- 1.2 Analyze design principles of a WLAN deployment
- 1.2.a Wireless deployment models (centralized, distributed, controller-less, controller based, cloud, remote branch)
- 1.2.b Location services in a WLAN design

This chapter takes a look at the various wireless LAN (WLAN) deployment models. Network engineers have numerous choices when designing a WLAN using wireless LAN controllers (WLCs) with different capabilities and form factors. They need to decide whether to use an on-premises or cloud-based management controller or even a controller-less deployment. This chapter examines WLAN deploying models, including autonomous, centralized, Cisco FlexConnect, Software-Defined Access (SD-Access), and cloud-based models. This chapter also looks at how Cisco Connected Mobile Experiences (CMX) provides location services to track client movement and provide analytic data.

**This chapter covers the following technology topics:**

- Wireless Deployment Models
  - Autonomous
  - Centralized
  - Cisco FlexConnect
  - Cloud-based
  - Embedded
- Wireless Location Services

## Cram Saver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. Which wireless deployment model has an access point operating in lightweight mode?
2. Which wireless deployment model has the WLC integrated into the access layer switch?

### Answers

1. Centralized mode
2. Embedded mode

# Wireless Deployment Models

### ExamAlert

Before taking the ENCOR exam, make sure you completely understand the various wireless deployment models.

This section covers several wireless deployment models:

- **Autonomous:** This type of deployment is used in smaller offices and hotspots. The access points (APs) are managed individually rather than through a controller. The autonomous deployment model provides a cost-effective and straightforward way to deploy wireless solutions.
- **Centralized:** With this type of deployment, the APs operate in lightweight mode rather than in autonomous mode. In lightweight mode, the AP loses its self-sufficiency and instead is joined to a WLC to become fully functional.
- **Cisco FlexConnect:** This type of deployment allows you to deploy wireless solutions to branch and remote sites without deploying controllers in each office. It is generally used in sites with a small number of APs where a WLC deployment cannot be justified or is not desired.

- **Software-Defined Access (SD-Access):** This type of deployment integrates wireless access into the SD-Access architecture to realize all the advantages of fabric and Cisco Digital Network Architecture (DNA) Center automation. In SD-Access wireless, the control plane is centralized, and the data plane is distributed using Virtual Extensible LAN (VXLAN) directly from fabric-enabled APs. Chapter 23, "SD-Access," covers SD-Access wireless, along with the other components of an SD-Access infrastructure, in more detail.
- **Cloud-based:** This type of virtual wireless controller is offered by Cisco as part of Cisco Meraki or Cisco Catalyst 9800 wireless solutions. This type of deployment offers centralized management and can scale from small to large networks.
- **Embedded:** In this deployment model, the WLC is co-located with the access layer switch. It is an embedded wireless topology in the sense that the WLC is embedded into the switch hardware. It can be desirable when the WLC function is embedded within the switch platform because you have fewer devices to manage and can potentially reduce operating costs. There is also a variation of the embedded model in which the WLC feature is embedded into Cisco Catalyst APs.

The following sections examine these wireless deployment models in more detail.

# Autonomous Wireless Deployments

Cisco APs can be configured to operate in either autonomous mode or lightweight mode. In autonomous mode, also known as standalone mode, the APs are self-contained, with each offering one or more fully functional basic service sets (BSSs). The APs are basically an extension of the switched network, connecting service set identifiers (SSIDs) to VLANs at the access layer. Each AP operates independently of the others and has no idea about other standalone APs. For this reason and because autonomous APs cannot scale to large environments, autonomous APs should not be deployed in business-critical environments.

However, deploying APs in autonomous mode has advantages. It is a simple and cost-effective way of extending the wired network in a single location or a small environment. If Wi-Fi coverage is needed in a few rooms or at a small office, a few APs can be deployed at the access layer in autonomous mode.

For management, each AP is configured with a management IP address that is be managed through Telnet, SSH, or a web interface. The Cisco IOS CLI can also be managed through the console port of the AP. Each AP deployed at a location needs to be managed individually. Typically, you deploy a dedicated management VLAN where the AP management IP will sit. Both the management and data VLANs need to be added to the trunk link to reach the AP.

# Centralized Wireless Deployments

As already discussed in this chapter, an AP can be configured to operate in autonomous mode or lightweight mode. In lightweight mode, the AP loses its self-sufficient ability to provide a BSS for wireless users. Instead, an AP joins a WLC to become fully functional in what is known as a *split-MAC architecture*, where the AP handles the real-time 802.11 processes, and the WLC performs management functions. The join process between the AP and WLC is done over a logical pair of Control and Provisioning of Wireless Access Points (CAPWAP) tunnels that extend through the wired network infrastructure. Control and data traffic are carried over these tunnels.

The Internet Engineering Task Force (IETF) CAPWAP standard is the underlying protocol used in the Cisco centralized WLAN architecture, which is the functional architecture of the Cisco Unified Wireless Network solution. CAPWAP enables the configuration and management of APs and WLANs in addition to encapsulation and forwarding of WLAN client traffic between an AP and a WLC.

CAPWAP is based on Lightweight Access Point Protocol (LWAPP) and adds additional security with Datagram Transport Layer Security (DTLS). CAPWAP uses User Datagram Protocol (UDP) and can operate over either Internet Protocol Version 4 (IPv4) or Internet Protocol Version 6 (IPv6).

IPv6 mandates a complete payload checksum for UDP, which impacts the performance of the AP and the WLC. To maximize performance for IPv6 deployments, the AP and WLC implement UDP Lite, which performs a checksum on only the header rather than on the full payload.

The key features of CAPWAP include split-MAC architecture, encryption, Layer 2 tunnels, and WLC discovery and selection. The split-MAC architecture is worth highlighting here.

The split-MAC concept is where the CAPWAP AP manages part of the 802.11 protocol operation, and the WLC manages the remaining aspects. With split-MAC architecture, simple, timing-dependent operations are generally managed

locally on the CAPWAP AP, and more complex, less time-dependent operations are managed on the WLC.

The CAPWAP AP handles the following:

- The frame exchange handshake between a client and an AP
- Transmission of beacon frames
- Buffering and transmission of frames for clients in power-save mode
- Responses to probe request frames from clients; the probe requests are also sent to the WLC for processing
- The forwarding notification of probe requests received by the WLC
- Provision of real-time signal quality information to the switch with every received frame
- Monitoring of each of the radio channels for noise, interference, and other WLANs
- Monitoring for the presence of other APs
- Encryption and decryption of 802.11 frames

A WLC provides the following MAC layer functions:

- 802.11 authentication
- 802.11 association and re-association
- 802.11 frame translation and bridging
- 802.1X/EAP/RADIUS processing
- Termination of 802.11 traffic on a wired interface, except in the case of FlexConnect APs

A CAPWAP tunnel supports two categories of traffic:

- **CAPWAP control messages:** These messages are used to convey control, configuration, and management information between the WLC and APs.
- **Wireless client data encapsulation:** Encapsulation is used to transport Layer 2 wireless client traffic in IP EtherType encapsulated packets from the AP to the WLC.

When the encapsulated client traffic reaches the WLC, it is mapped to a corresponding interface (VLAN) or interface group (VLAN pool) at the WLC. This interface mapping is defined as part of the WLAN configuration settings on the WLC. The interface mapping is usually static; however, a WLAN client may also be dynamically mapped to a specific VLAN, depending on the local policies defined on the WLC or on RADIUS return attributes forwarded from an upstream AAA server upon successful authentication.

Encryption is provided for CAPWAP control and data packets exchanged between an AP and a WLC by using DTLS. DTLS is an IETF protocol based on TLS. All Cisco access points and controllers are shipped with a manufacturer installed certificate (MIC), which an AP and WLC use by default for mutual authentication and encryption key generation.

DTLS is enabled by default to secure the CAPWAP control channel but is disabled by default for the data channel. All CAPWAP management and control traffic exchanged between an AP and a WLC is encrypted and secured by default to provide control plane privacy and prevent man-in-the-middle attacks. CAPWAP data encryption is optional and is enabled for each AP. When enabled, all WLAN client traffic is encrypted at the AP before being forwarded to the WLC and vice versa. Enabling DTLS data encryption impacts the performance of both the APs and the WLCs. DTLS data encryption should only be enabled on APs deployed over an unsecured network.

Unlike LWAPP, which operates in either a Layer 2 or a Layer 3 mode, CAPWAP only operates in Layer 3 and requires IP addresses to be present on both the AP and the WLC. CAPWAP uses UDP for IPv4 deployments and UDP or UDP Lite (the default) for IPv6 implementations to facilitate communication between an AP and a WLC over an intermediate network. CAPWAP can perform fragmentation and reassembly of tunnel packets, allowing WLAN client traffic to make use of a full 1500-byte MTU without having to adjust for any tunnel overhead.

A lightweight AP discovers a WLC by using a CAPWAP discovery mechanism and then sends the controller a CAPWAP join request. When an AP joins a WLC, the WLC manages the AP configuration, firmware, control transactions, and data transactions. A CAPWAP AP must discover and join a WLC before becoming an active part of the Cisco Unified Wireless Network.

Figure 20.1 shows a centralized deployment with guest wireless. Here, CAPWAP control messages and wireless client data moved from the Catalyst 9100 AP over to the Catalyst 9800 WLC. Guest wireless traffic also moves over the CAPWAP tunnel.

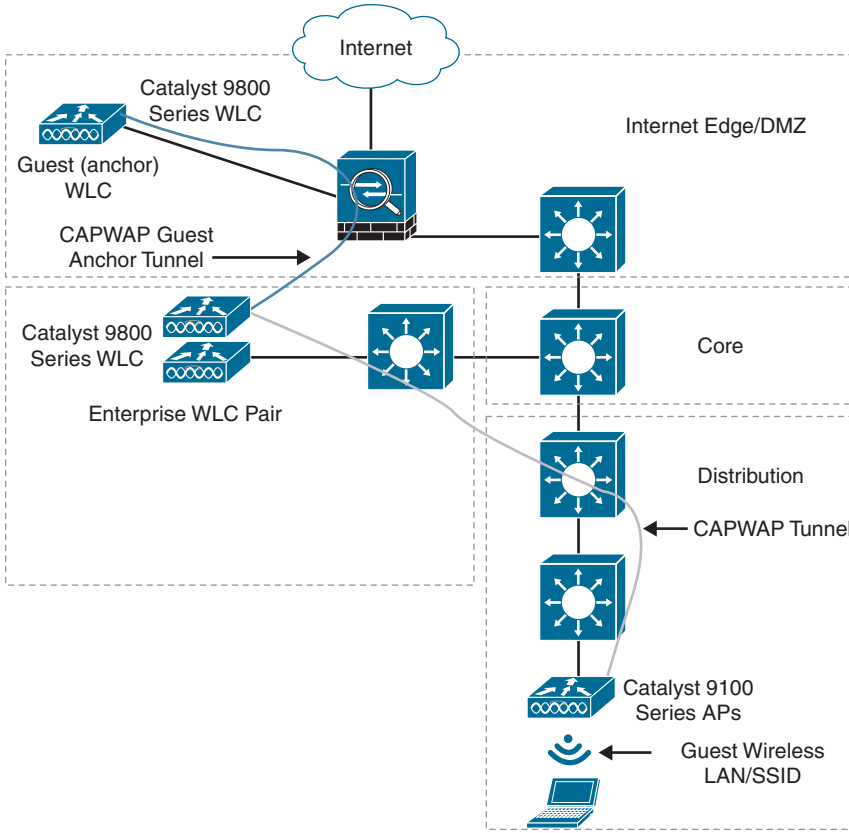



FIGURE 20.1 **Centralized Deployment with Guest Wireless**

# Cisco FlexConnect Wireless Deployments

The Cisco FlexConnect wireless solution is designed for branch office and remote office deployments. It allows for the configuration and control of APs in a branch or remote office from the corporate office through a WAN link without requiring controller deployment in each office. The FlexConnect APs can switch client data traffic locally and perform client authentication locally. When they are connected to the controller, they can also send traffic back to the controller.

A FlexConnect deployment allows enterprises to:

- Centralize control and manage the traffic of APs from the main site
- Distribute the client data traffic at each branch site

Both the centralized and the distributed models have advantages. These are the advantages of centralizing APs control traffic:

- Single pane of monitoring and troubleshooting
- Ease of management
- Secured and seamless mobile access to data center resources
- Reduction in branch footprint
- Increase in operational savings

These are the advantages of distributing client data traffic:

- No operational downtime (survivability) against complete WAN link failures or controller unavailability
- Mobility resiliency within the branch during WAN link failures
- Increase in branch scalability (up to 100 APs and 250,000 square feet per branch and 5000 square feet per AP)

There are two modes of operation for FlexConnect APs:

- **Connected mode:** The WLC is reachable. In this mode, the FlexConnect AP has CAPWAP connectivity with its WLC.
- **Standalone mode:** The WLC is unreachable. The FlexConnect AP has lost or failed to establish CAPWAP connectivity with its WLC (for example, when there is a WAN link outage between a branch and its main site).

Figure 20.2 shows a Cisco FlexConnect deployment with guest wireless in connected mode. When there is connectivity between the AP and the WLC, traffic flows through the CAPWAP tunnel (in connected mode). In the absence of connectivity between the AP and WLC, the FlexConnect APs can switch client data traffic locally and perform client authentication locally.

Figure 20.3 shows a Cisco FlexConnect deployment with guest wireless in standalone mode. When a FlexConnect access point cannot access the WLC, the access point enters standalone mode and authenticates clients by itself. When a client associates to a FlexConnect access point, the access point sends all authentication messages to the controller and either switches the client data packets locally (locally switched) or sends them to the controller (centrally switched), depending on the WLAN configuration.

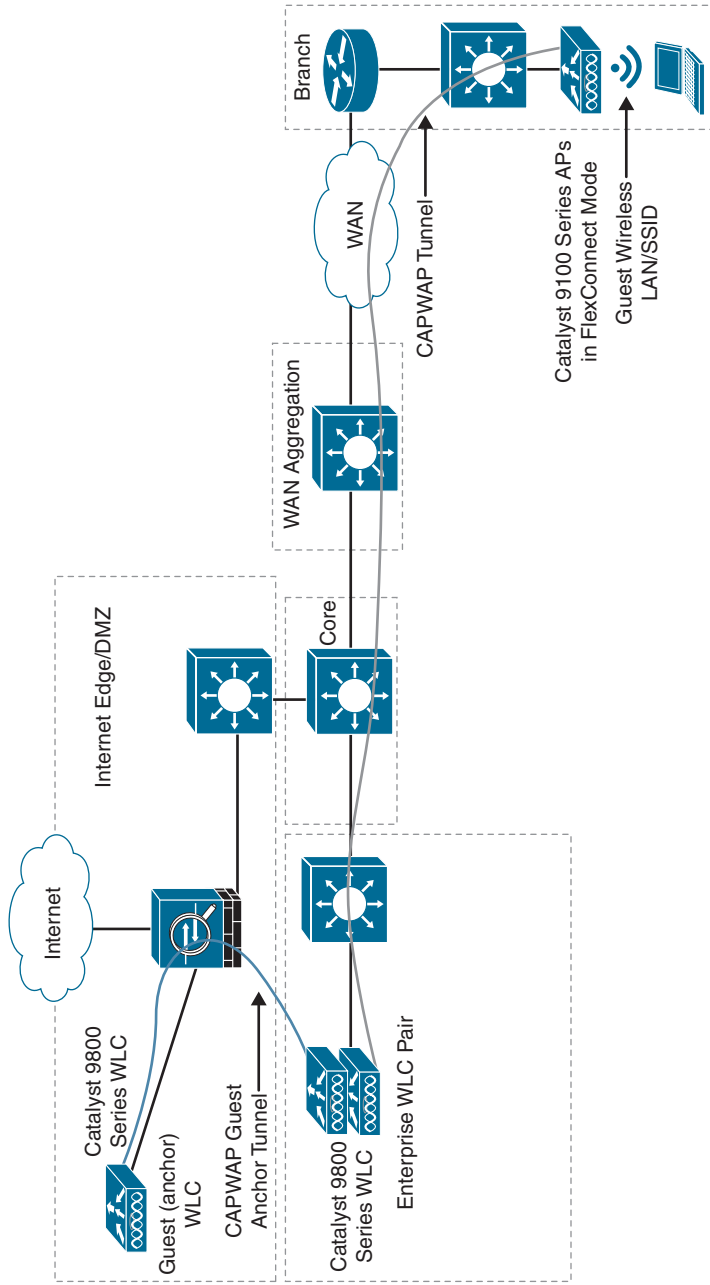



FIGURE 20.2 **Cisco FlexConnect Deployment with Guest Wireless in Connected Mode**

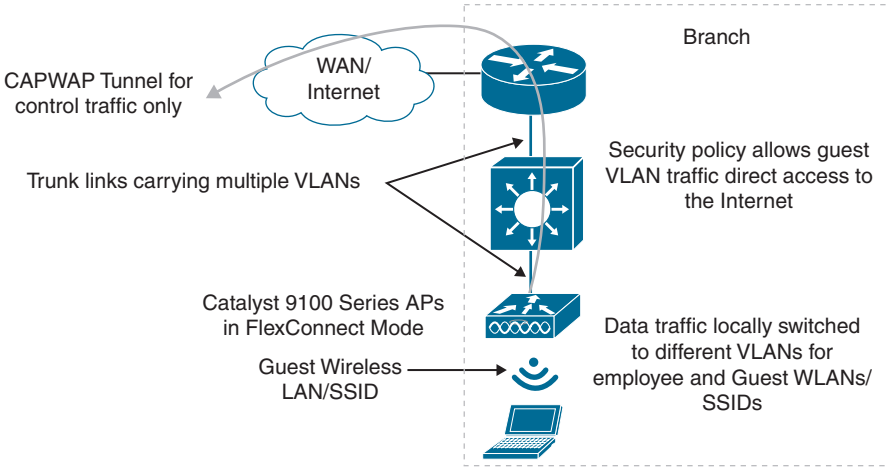



FIGURE 20.3 **Cisco FlexConnect Deployment with Guest Wireless in Standalone Mode**

With FlexConnect, when an AP boots up, it looks for a controller. If it finds one, it joins the controller, downloads the latest software image and configuration from the controller, and initializes the radio. It saves the downloaded configuration in nonvolatile memory for use in standalone mode. Once the access point is rebooted after the latest controller software is downloaded, it must be converted to FlexConnect mode. This can be done using the GUI or the CLI.

When a FlexConnect access point can reach the controller (referred to as the *connected mode*), the controller assists in client authentication. When a FlexConnect access point cannot access the controller, the access point enters standalone mode and authenticates clients by itself. When a client associates to a FlexConnect access point, the access point sends all authentication messages to the controller and either switches the client data packets locally (locally switched) or sends them to the controller (centrally switched), depending on the WLAN configuration.

# Cloud-Based Wireless Deployments

WLCs no longer need to be a physical on-premises appliances. The same controller functionality can easily be deployed as software in the public cloud. The Cisco Meraki portfolio gives you the flexibility to use a web-based management platform hosted off-premises to manage APs. You can also use the Cisco Catalyst 9800 Series WLC, which is built for intent-based networking. The controller functionality can be hosted in either a private cloud or a public cloud.

Cisco's Meraki cloud management platform provides visibility to network users, their devices, and their applications. With a rich set of analytics, administrators can quickly create access controls and application usage policies, optimizing both the end-user experience and network security. In addition to the visibility that the platform gives, it allows for multisite management, which eliminates most of the costs and complexities associated with an on-premises WLC.

The Meraki platform also provides for zero-touch provisioning, which shortens deployment and configuration time to minutes. There is no need to manually stage APs or perform manual configuration and provisioning. Meraki APs uses Auto RF to self-configure and optimize RF settings for maximum performance, even in dense and challenging environments.

Another plus of the cloud-based Meraki platform is its ability to seamlessly provide over-the-Web upgrades for APs and Meraki switches that support those APs. New feature updates are added automatically to the Meraki platform to provide improved visibility, analytics, security, and troubleshooting tools to enhance administrator efficiency.

The data (configuration, statistics monitoring, and so on) that flows from APs to the Meraki cloud flows over a secure link. User data does not flow through the Meraki cloud. Instead, it flows directly to its destination on the LAN or across the WAN. Meraki uses an event-driven remote procedure call (RPC) engine for Meraki devices to communicate to the Meraki dashboard and for Meraki servers to send and receive data. The Meraki APs act as servers/receivers, and the Meraki cloud initiates calls to the devices for data collection and configuration deployment. Because the cloud infrastructure is the initiator, configurations can be executed in the cloud before the devices are actually online or even physically deployed.

In the event of cloud connectivity loss (for example, caused by a local ISP or connection failure), the APs continue to run with their last known configuration until cloud connectivity is restored. However, if an AP goes offline due to a break in connectivity, it attempts to communicate with the controller in the cloud until it regains connectivity. Once the AP comes online, it automatically receives the most recent configuration settings from the Meraki cloud. If changes are made to the device configuration while the device is online, the device automatically receives and updates these changes. These changes are generally available on the device in a matter of seconds. If the administrator makes no configuration changes, the device continues to periodically check for updates to its configuration on its own.

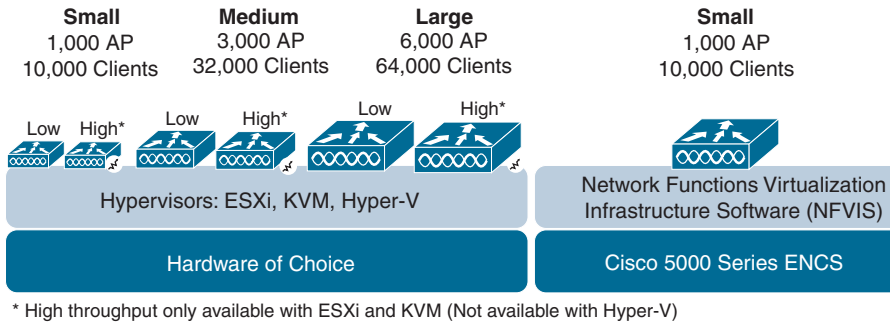
For devices to communicate with the cloud, Meraki leverages a proprietary lightweight encrypted tunnel using AES-256 encryption while management data is in transit. Within the tunnel itself, Meraki leverages HTTPS and Protocol Buffers for a secure and efficient solution that is limited to 1 Kbps per device when the device is not being actively managed.

The other option you have for cloud-based controller deployment is the Cisco Catalyst 9800-CL wireless controller. The Catalyst 9800-CL was built from the ground up for intent-based networks and Cisco DNA. It is based on Cisco IOS XE and integrates with Cisco Aironet APs.

The Catalyst 9800-CL can be deployed both in private and public cloud environments. Some of the key highlights of its operation in private cloud operations are as follows:

- It supports VMware ESXi, KVM, Hyper-V hypervisors, and Cisco NFVIS (on ENCS).
- It supports centralized, Cisco FlexConnect, mesh, and fabric (SD-Access) deployment modes.
- It supports multiple scales and throughput profiles with a single deployment package:
  - **Small (low/high throughput):** Designed for distributed branches and small campuses supporting up to 1000 APs and 10,000 clients
  - **Medium (low/high throughput):** Designed for medium-sized campuses supporting up to 3000 APs and 32,000 clients
  - **Large (low/high throughput):** Designed for large enterprises and service providers supporting up to 6000 APs and 64,000 clients
- It contains one deployment package for all the scale templates. You choose the deployment size and the throughput profile when instantiating the virtual machine (VM).
- It attains, with a high (enhanced) throughput profile, up to 5 Gbps on ESXi and KVM with the right set of network cards and resources (such as an SR-IOV-enabled NIC); the high-throughput profile is not available in a Hyper-V environment.
- It provides an intuitive bootstrap wizard for VM instantiation to boot the wireless controller with recommended parameters.
- It works ideally for branch deployments. The Catalyst 9800-CL can be deployed as a virtual machine on the Cisco 5000 Series Enterprise Network Compute System (ENCS) running Cisco NFVIS at branch sites.

Figure 20.4 shows the Cisco Catalyst 9800-CL for a private cloud.




* High throughput only available with ESXi and KVM (Not available with Hyper-V)

FIGURE 20.4 **Cisco Catalyst 9800-CL for a Private Cloud**

At this writing, the Catalyst 9800-CL can be deployed in two public cloud platforms: Amazon Web Services (AWS) and Google Cloud Platform (GCP).

Some of the critical highlights of Catalyst 9800-CL in public cloud operations are as follows:

- The Cisco Catalyst 9800-CL is available as an infrastructure-as-a-service (IaaS) solution on AWS and GCP Marketplaces.
- It is supported only with managed VPN deployment mode:
  - The 9800-CL should be instantiated within a virtual private cloud (VPC).
  - A VPN tunnel must be established from the customer site to AWS or GCP to enable communication between the Cisco AP and 9800-CL wireless controller.
- It supports up to 6000 access points and 64,000 clients.
- The wireless controller instance can be deployed in AWS using cloudformation templates provided by Cisco or manually using the EC2 console.
- The wireless controller can be deployed in GCP using the guided workflow in the marketplace.

Figure 20.5 shows the Cisco Catalyst 9800-CL for a public cloud.

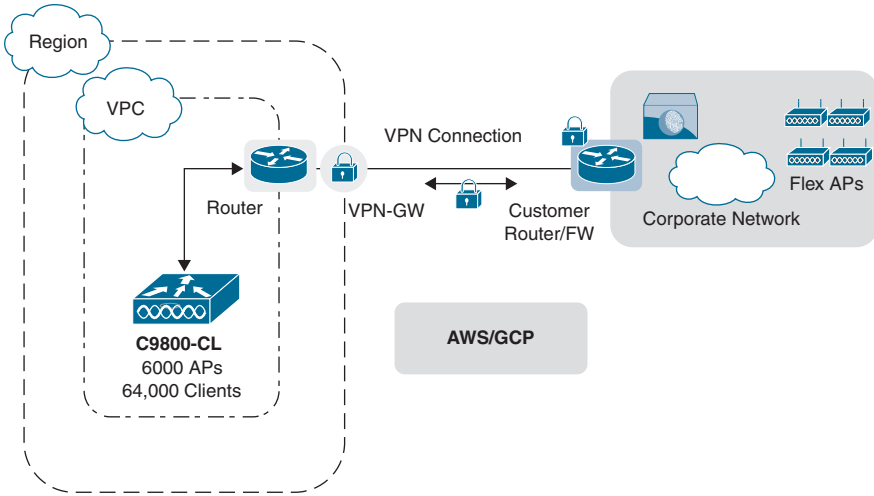



FIGURE 20.5 **Cisco Catalyst 9800-CL for a Public Cloud**

Whereas the Catalyst 9800-CL is designed for private and public environments, the Catalyst 9800 wireless platform can be deployed as the following:

- Physical WLC appliances for small to large campus deployments (centralized deployment model)
- Embedded wireless controller for switches
- Embedded wireless controller for APs

The centralized deployment model was covered earlier in this chapter, and the embedded options are covered next.

# Embedded Wireless Deployments

For small to medium-sized branch wireless network deployments, it may be feasible to deploy the Cisco Embedded Wireless Controller (EWC) on Cisco Catalyst 9300, 9400, and 9500 switch platforms. (A non-SD-Access deployment option is covered later in this section.) Cisco EWC combines these Catalyst 9000 Series switches with the Catalyst 9800 Series wireless controllers in a single box.

In some branch sites, you might have heavily invested in Cisco Catalyst platforms already, and integrating wireless can help to maximize the value of such a platform. Wireless capabilities can be easily enabled without the need for an additional standalone controller. Once EWC is deployed, small to

medium-sized environments can simply connect to the switch's WebUI to deploy corporate and guest WLANs.

To use EWC on a switch, the following requirements need to be met:

- **Catalyst 9000 switch family:** The switch platform needs to be Catalyst 9300, 9300L, 9500, or 9500H Series.
- **IOS XE:** The wireless subpackage must be installed.
- **Management IP address:** This is required to access the WebUI.
- **APs:** These need to be Catalyst 9100 or Aironet 1800, 2800, 3800, 4800, 1540, or 1560 Series.
- **The switches and APs:** Both require Cisco DNA Advantage licenses.

EWC can be deployed in several ways:

- With a single switch
- With High Availability Stateful Switchover (SSO)
- With N+1 redundancy

Another embedded option is Cisco Embedded Wireless Controller on Catalyst Access Points (EWC-AP). EWC-AP is a next-generation Wi-Fi solution that combines the Cisco Catalyst 9800 wireless controller with the latest Wi-Fi 6 Cisco Catalyst 9100 APs.

There is also a third embedded option, Cisco Mobility Express, which is covered shortly.

Figure 20.6 shows the various supported Cisco Catalyst 9100 Series APs for different deployments.


FIGURE 20.6 **Cisco Catalyst 9100 Series APs for Different Deployments**

Built for intent-based networking and Cisco DNA, the EWC-AP helps simplify complexity and reduce operational costs by leveraging intelligence and automation. Pairing the EWC-AP with Cisco DNA allows you to transform a network. Cisco DNA allows you to understand a network with real-time

analytics, quickly detect and contain security threats, and easily provide networkwide consistency through automation and virtualization.

EWC-AP can be managed through its web-based dashboard or through a mobile app (which can be used for deployment, provisioning, and monitoring). Also, you can use open-standards-based programmability with NETCONF and YANG. EWC-AP can be used for larger environments in a distributed office deployment where there is a need to manage multiple EWC networks from a centralized manager. This enables you to use Cisco DNA Center for automation and assurance. For this use case, a Cisco DNA subscription license and smart licensing are required.

With Cisco Mobility Express, you can deploy a high-performance 802.11ac Wave 2 (Wi-Fi 5) network. Cisco Mobility Express puts the wireless controller functionality into Cisco Aironet APs. It is a virtual WLC integrated on 802.11ac Wave 2 APs (Cisco Aironet 4800, 3800, 2800, 1850, 1830, 1815, 1560, and 1540 Series). Mobility Express is for small and medium-sized businesses, as well as distributed enterprises serving up to 100 access points and 2000 clients. Multiple sites can be deployed with Mobility Express and managed using Cisco DNA Center.

There is also support for redundancy with Cisco Mobility Express. If an access point running the wireless LAN controller function goes down for any reason, another Mobility Express access point is elected to run the wireless LAN controller function. To achieve redundancy, you need to deploy at least two Mobility Express APs in a network.

The AP acting as the WLC is referred to as the *primary AP*, and the other APs in the Cisco Mobility Express network, which are managed by the primary AP, are referred to as *subordinate APs*. In addition to acting as a WLC, the primary AP operates as an AP to serve clients along with the subordinate APs.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which of the following are advantages of the autonomous wireless deployment model? (Choose two.)
   - ❍ **A.** It is simple to deploy.
   - ❍ **B.** It can scale to large environments.
   - ❍ **C.** It is cost-effective.
   - ❍ **D.** It can be deployed with a WLC.

**2.** With the split-MAC architecture, association and re-association are handled by which device?

- ❍ **A.** Lightweight AP
- ❍ **B.** Autonomous AP
- ❍ **C.** WLC
- ❍ **D.** Access layer switch

**3.** With a Cisco Mobility Express solution, what type of AP is required to have the wireless controller function embedded in it?

- ❍ **A.** 802.11ac Wave 1 series APs
- ❍ **B.** 802.11ac Wave 2 series APs
- ❍ **C.** Cisco Catalyst 9100s series APs
- ❍ **D.** Any Cisco AP

**4.** Which type of wireless deployment model is most suitable for small offices and hotspots?

- ❍ **A.** Autonomous
- ❍ **B.** Centralized
- ❍ **C.** FlexConnect
- ❍ **D.** SD-Access

**5.** Virtual wireless controllers offered by Cisco Meraki and Cisco Catalyst 9800 Series wireless solutions are part of which wireless deployment model?

- ❍ **A.** Embedded
- ❍ **B.** Cloud-based
- ❍ **C.** FlexConnect
- ❍ **D.** SD-Access

**6.** Which wireless deployment model allows for the deployment of wireless solutions to branch and remote sites without requiring deployment of a wireless controller at the branch or remote site?

- ❍ **A.** Embedded
- ❍ **B.** Cloud-based
- ❍ **C.** FlexConnect
- ❍ **D.** SD-Access

# Answers

1. **A** and **C** are correct. Deploying APs in autonomous mode is a simple and cost-effective way of extending a wired network in a single location or a small environment.
2. **C** is correct. In a split-MAC architecture, the WLC handles association and re-association.
3. **B** is correct. Cisco Mobility Express puts the wireless controller functionality into Cisco Aironet APs and requires 802.11ac Wave 2 (Wi-Fi 5) APs.
4. **A** is correct. The autonomous wireless deployment model is suitable for smaller offices and hotspots.
5. **B** is correct. Cisco offers virtual wireless controllers as part of Cisco Meraki or Cisco Catalyst 9800 wireless solutions; they are part of Cisco's cloud-based offerings.
6. **C** is correct. Cisco FlexConnect wireless deployment allows you to deploy wireless solutions to branch and remote sites without deploying controllers in each office.

# Wireless Location Services

Research and development in Wi-Fi location prediction techniques have given rise to indoor RF location-tracking systems. Using location-based services helps you perform network analytics and facilitates the collection of meaningful data about the wireless part of a network. At the heart of generating better insights into wireless customer behavior is the Cisco Mobile Experience (CMX) solution.

CMX is a software solution that uses location and other intelligence from Cisco wireless infrastructure to generate analytics and deliver relevant services to customers on their mobile devices. With CMX, organizations can easily onboard users to a wireless network directly, serve personalized content to users on their mobile devices, enhance the in-venue experience, and generate better insights into customer behavior and venue space utilization.

In a nutshell, CMX helps create personalized mobile experiences for end users and provides operational efficiency with location-based services. It helps customers determine the locations of devices in their network and can be used for various location-based services.

CMX transforms a network into a customer experience engine. CMX enables organizations to detect, connect, and engage with end users while inside their venue:

- **Detect:** The wireless signal from the customer's mobile device and its characteristics are detected as the customer approaches the location.
- **Connect:** After the customer receives notification of available Wi-Fi access and services, the customer can securely connect.
- **Engage:** When customers have access, you can engage them with personalized content. Through two-way communication, you can build real-time, value-added relationships with customers.

The capabilities of CMX are delivered via three components:

- Location
- CMX Connect
- CMX Analytics

CMX uses existing wireless infrastructure to calculate the location of the Wi-Fi devices and interferers (for example, microwave ovens) in the network. It offers location capabilities from proximity-based (presence) to highly accurate coordinates on a map (hyperlocation).

CMX Connect delivers targeted, context-specific experiences to on-site visitors. It provides an easy way to create customizable captive portals and capture visitor information through multiple onboarding options. Through the data collected, CMX Connect allows organizations to engage with visitors on the captive portal or through external media, such as mobile applications, digital signage, or offline marketing.

CMX Analytics generates insights into the Wi-Fi devices of visitors in the venue based on their location and movement patterns. There are two flavors of analytics:

- **Presence analytics:** CMX Analytics uses proximity information from Wi-Fi devices to generate real-time and historical insights into visitor and passerby behaviors. Proximity is determined by the signal strength and time duration of Wi-Fi devices detected by the nearest access points. Presence Analytics captures metrics such as device counts, dwell times, and repeat visitors' breakdowns.
- **Location analytics:** CMX Analytics uses x,y coordinates calculated by the CMX Location engine to provide real-time and historical location data at a more granular level and to subsequently generate better insights.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. What CMX features allow organizations to interact with end users? (Choose three.)
   - ❍ **A.** Detect
   - ❍ **B.** Presence
   - ❍ **C.** Engage
   - ❍ **D.** Connect
   - ❍ **E.** Location

2. What are the flavors of analytics provided by CMX? (Choose two.)
   - ❍ **A.** Detect
   - ❍ **B.** Presence
   - ❍ **C.** Engage
   - ❍ **D.** Location
   - ❍ **E.** Connect

3. Which type of analytics captures metrics such as device counts, dwell times, and repeat visitors' breakdowns?
   - ❍ **A.** Detect
   - ❍ **B.** Presence
   - ❍ **C.** Engage
   - ❍ **D.** Location
   - ❍ **E.** Connect

# Answers

1. **A**, **C**, and **D** are correct. CMX enables organizations to detect, engage, and connect with end users while inside their venue.
2. **B** and **D** are correct. CMX Analytics generates insights into the Wi-Fi devices of visitors in a venue based on their location and movement patterns. These are grouped as Presence Analytics and Location Analytics.
3. **B** is correct. Presence Analytics captures metrics such as device counts, dwell times, and repeat visitors' breakdowns.

# Review Questions

1. Deployment of autonomous APs is typically done in which one of the following situations?
   - ❍ **A.** Deployment at the main site
   - ❍ **B.** Deployment for small offices or rooms
   - ❍ **C.** Deployment for large branch sites
   - ❍ **D.** Cloud-based deployment

2. In a split-MAC architecture, the AP handles all except which of the following operations?
   - ❍ **A.** Frame exchange handshake between a client and an AP
   - ❍ **B.** Monitoring of each of the radio channels for noise, interference, and other WLANs
   - ❍ **C.** Encryption and decryption of 802.11 frames
   - ❍ **D.** 802.1X/EAP/RADIUS processing

3. The Catalyst 9800-CL can be deployed in all except which of the following private cloud environments?
   - ❍ **A.** Cisco NFVIS
   - ❍ **B.** ESXi
   - ❍ **C.** VMware Fusion
   - ❍ **D.** KVM

4. For a non-SD-Access Cisco Embedded Wireless Controller (EWC) deployment, all except which of the following requirements must be met?
   - ❍ **A.** The switch platform needs to be Catalyst 9300, 9300L, 9500, or 9500H Series.
   - ❍ **B.** The wireless sub-package is installed on IOS XE.
   - ❍ **C.** The switches and APs require Cisco DNA Advantage licenses.
   - ❍ **D.** DNA Center must be deployed in the environment.

5. Which Cisco embedded wireless solution embeds the Cisco Catalyst 9800 wireless controller function into the latest Wi-Fi 6 Cisco Catalyst 9100 APs without the need for a dedicated WLC?
   - ❍ **A.** EWC-AP
   - ❍ **B.** Cisco Mobility Express
   - ❍ **C.** EWC
   - ❍ **D.** Catalyst 9800-CL

# Answers to Review Questions

1. **B** is correct. Autonomous APs are typically deployed in smaller offices and hotspots.
2. **D** is correct. The WLC does 802.1X/EAP/RADIUS processing.
3. **C** is correct. VMware ESXi, KVM, Hyper-V hypervisors, and Cisco NFVIS (on ENCS) are all supported private cloud platforms for deploying Cisco Catalyst 9800-CL wireless controllers.
4. **D** is correct. For a non-SD-Access Cisco Embedded Wireless Controller (EWC) deployment, DNA Center is not a requirement.
5. **A** is correct. Cisco Embedded Wireless Controller on Catalyst Access Points (EWC-AP) combines the Cisco Catalyst 9800 wireless controller function into the latest Wi-Fi 6 Cisco Catalyst 9100 APs. There is no need for a dedicated WLC.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers on-premises vs. cloud infrastructure.

*This page intentionally left blank*

CHAPTER 21

# On-Premises vs. Cloud Infrastructure

**This chapter covers the following official ENCOR 350-401 exam objective:**

- 1.3 Differentiate between on-premises and cloud infrastructure deployments

Over the past decade, cloud computing has been one of the most exciting and game-changing technologies in the tech marketplace, and its use has expanded greatly. Cloud computing is an approach to a data center infrastructure that allows several services to be available to several users in an environment. The cloud infrastructure is generally part of a delivery model that determines how the cloud infrastructure will be used, who operates the cloud, and where the data is located. This chapter looks at the four most common cloud delivery or deployment models: public cloud, private cloud, hybrid cloud, and community cloud. This chapter also covers the common cloud services models: infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS).

**This chapter covers the following technology topics:**

- Cloud Infrastructure Basics
- Cloud Services Models
- Cloud Deployment Models
- On-Premises or Cloud Infrastructure

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. In the delivery of cloud services, what is the meaning of *resource pooling*?
2. Which cloud services model allows a cloud services consumer to use arbitrary software, including operating systems and applications?
3. Which cloud deployment or delivery model is a composition of two or more distinct cloud infrastructures?
4. What is one primary consideration for keeping workload on-premises?

### Answers

1. With resource pooling, a service provider's computing resources are pooled together in a multitenancy model to serve multiple consumers.
2. Infrastructure as a service (IaaS)
3. Hybrid cloud
4. Keeping workload on-premises provides a sense of being more secure because applications and data are on an organization's own server, in its own storage, and behind its own firewalls.

# Cloud Infrastructure Basics

The U.S. National Institute of Standards and Technology (NIST) describes cloud computing as follows:

> Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model comprises five essential characteristics, three service models, and four deployment models.

For services to be considered cloud services, they must meet certain essential characteristics. NIST defines these essential characteristics as follows:

- **On-demand self-service:** The cloud services consumer can unilaterally provision the cloud capabilities automatically through automation

and orchestration tools provided by the cloud services provider. The idea here is that the requests for the use of computing, network, and storage resources are self-served and do not require human interaction from the cloud services provider side. It is essential that cloud users be able to perform self-service so that they can quickly provision service instances as the demand arises.

- **Broad network access:** The cloud capabilities are accessed through a standard mechanism that uses thin or thick client platforms. Cloud services consumers can access the cloud using mobile devices, laptops, or workstations.
- **Resource pooling:** The service provider's computing resources are pooled together in a multitenancy model to serve multiple consumers. Based on the consumers' demands, the virtual resources can be dynamically assigned and reassigned. Generally, consumers cannot control the exact resource locations where their workloads run. However, they can control the resource's location at a high level (for example, the country, state, data center, or availability zone).
- **Rapid elasticity:** Cloud resources can be automatically and rapidly expanded and contracted to meet the demand for the resources. From the consumer perspective, it appears as if the cloud services provider has an unlimited quantity of resources. It is essential for the cloud services provider to have resources elastically sized to meet the cloud consumer's changing demands. It should be noted here that some cloud services providers may still need a formal request to allow expansion beyond certain fixed quotas that they may have set by default to prevent unintentionally high usage by consumers.
- **Measured service:** Transparency is provided for both the cloud services provider and the cloud services consumer, so there needs to be some sort of metering capability. The metering capability monitors, controls, and reports on the storage, compute, and network bandwidth usage for a particular user. Metering is usually done on a per-user basis. A cloud services provider can offer low competitive pricing due to economies of scale. The consumer can access cloud services with little to no capital investment in infrastructure setup.

Figure 21.1 shows the essential characteristics of a cloud infrastructure.

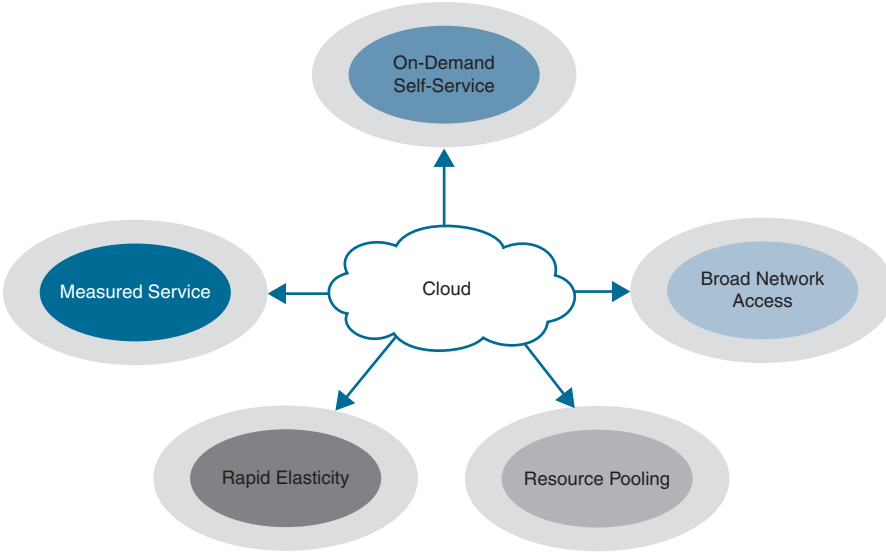



FIGURE 21.1 **Characteristics of a Cloud Infrastructure**

The infrastructure that facilitates these key characteristics of cloud services depends on a couple key enablers. These enablers allow cloud services providers to provision cloud services to large- and small-scale consumers. Although these enablers are not part of the definitions of cloud computing from NIST, they are a critical part of the cloud services provider infrastructure that allows for cloud services provisioning.

These are the key enablers:

- **Virtualization:** Virtualization of computing and network resources allows for the pooling and allocation of resources on demand. Virtualization also allows for metering capabilities that facilitate billing for resource usage. (Virtualization is covered in Chapter 26, "Basic Virtualization.")
- **Automation:** Automation facilitates the elastic use of cloud resources and allows for workload mobility on the underlying resources.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which of the following is not a characteristics of cloud services, as identified by NIST?
   - ❍ **A.** Measured service
   - ❍ **B.** Rapid elasticity
   - ❍ **C.** Customer-provided licenses
   - ❍ **D.** On-demand self-service

2. True or false: The metering service of most cloud services providers provides monitoring and reporting on the storage, compute, and network bandwidth usage for a particular user.
   - ❍ **A.** True
   - ❍ **B.** False

## Answers

1. **C** is correct. Customer-provided licenses is not a characteristics of cloud computing; customers generally provide licenses when running on-premises infrastructure.
2. **A** is correct. The metering capability monitors, controls, and reports on the storage, compute, and network bandwidth usage for a particular user.

# Cloud Services Models

When discussing the various cloud services offered by cloud services providers in the market, it can help to group the service capabilities.

**ExamAlert**

Before taking the ENCOR exam, have a complete understanding of the various cloud services models.

NIST groups cloud services into service models. Each of the NIST service models provides a level of abstraction that reduces the effort needed to deploy cloud services on the cloud services provider infrastructure. NIST identifies three cloud services models: infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS).

## Infrastructure as a Service (IaaS)

With IaaS, a cloud services provider offers cloud services consumers the ability to provision compute, storage, and network resources. The cloud services consumers retain control and management of the operating systems, storage, and applications they deploy and are responsible for updating and securing their deployed systems. The underlying infrastructure environment remains under the control of the cloud services provider. The cloud services provider is responsible for hardware maintenance and virtualization platform maintenance, underlying networking, underlying storage, underlying infrastructure security, and potential backups of the cloud services consumer infrastructure.

IaaS provides cloud services users with the highest level of flexibility and control of the cloud services. The IaaS cloud services model is similar to what on-premises IT teams and developers are familiar with. They may be given resources to deploy operating systems and applications on hardware and are responsible for managing, securing, and updating them.

Rather than purchasing the physical hardware, storage, networking, software, and data center space with cooling, a cloud services consumer buys all of those in the cloud, for a recurring or fixed operating expense. Compared to on-premises infrastructure, with IaaS, the cost associated with maintaining physical infrastructure components (computing, storage, and network) is eliminated. It cuts out the customer's need to procure, ship, rack, configure, and test equipment.

With IaaS, the infrastructure is available on demand through the power of virtualization. The cloud services provider can provision resources on demand by using a web-based console or calling an application programming interface (API). Once the infrastructure is up and running, cloud services consumers are billed much as they are for utilities. Usage of the virtualized infrastructure is metered. Cloud services consumers accumulate systems costs when their infrastructure instances are powered on, and they stop accumulating costs when the instances are powered off.

IaaS provides an effective platform that is scalable, agile, and flexible. It speeds up time to market and frees the end user or business from managing data centers and infrastructure to focus more on building and managing systems applications.

Some popular IaaS offerings are Amazon Web Services (AWS) EC2, Google Cloud Platform (GCP), Microsoft Azure, and DigitalOcean.

Figure 21.2 shows the IaaS infrastructure with the cloud services provider and cloud services consumer layers of responsibility.

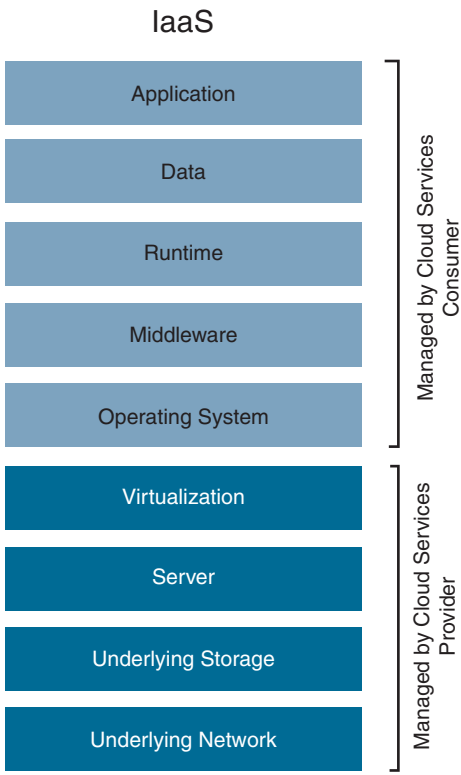



FIGURE 21.2 **IaaS Layers of Responsibility**

# Platform as a Service (PaaS)

With PaaS, the cloud services consumer can build on top of the cloud services provider infrastructure by using programming languages, libraries, services, and tools. The underlying infrastructure, including compute, network, and storage, remains under the cloud services provider's control. The cloud services consumer has control over the deployed application and may have some control over the application-hosting environment.

PaaS abstracts application-level functions and offers them as a service to the cloud services consumers. This way, developers can focus on the business needs rather than managing the underlying hosting infrastructure. The cloud services provider provides a suite of protocols to the cloud services consumer (in this case, the developer) to support development processes. However, the developer is limited or constrained by the cloud services provider's tools and software stack.

Some examples of PaaS are AWS Elastic Beanstalk, Google App Engine, and OpenShift.

Figure 21.3 shows the PaaS infrastructure with the cloud services provider and cloud services consumer layers of responsibility.

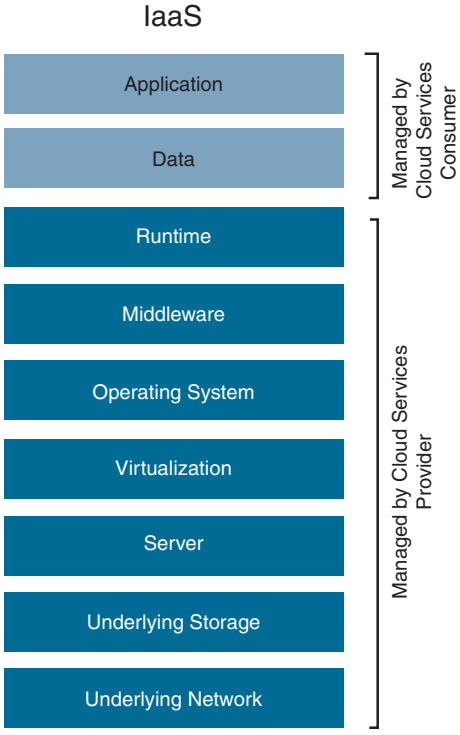



FIGURE 21.3 **PaaS Layers of Responsibility**

# Software as a Service (SaaS)

With SaaS, the cloud services provider makes available its applications for use by the cloud services consumer. The cloud services consumer does not control the underlying cloud infrastructure, including compute, network, and storage. The cloud services consumer can, however, manage and control limited user-specific application control settings. The SaaS applications are accessible through a client interface, which could be web based (for example, web-based email) or an application interface on a thin client or workstation (for example, an email application).

The SaaS model provides a complete application that is ready to use and is sometimes referred to as *on-demand software* or *hosted software*. It is a simplified model where it is common for the cloud services consumer to use a software subscription model. From the cloud services user's perspective, there is no complicated installation and software update, thus reducing the need for technical support. Users simply connect and start using a remote application hosted remotely on the cloud services provider infrastructure. Some common examples of the SaaS model are Dropbox, Google Workspace, GoToMeeting, and Cisco WebEx.

Figure 21.4 shows the SaaS infrastructure layers of responsibility.

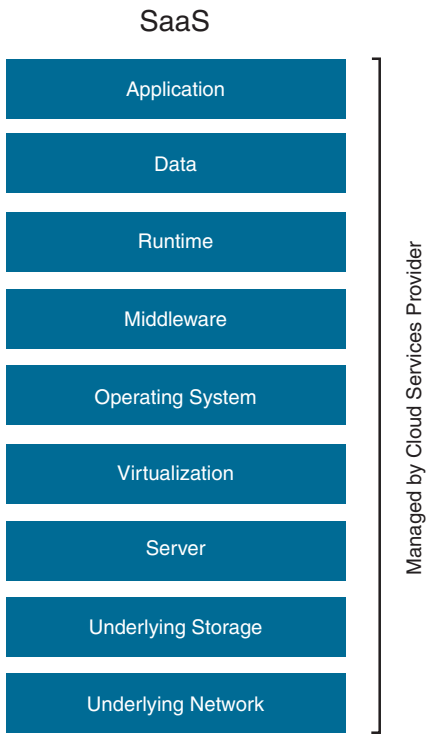



FIGURE 21.4 **SaaS Layers of Responsibility**

# Anything as a Service (XaaS)

In addition to the three common cloud services models just discussed, there is a broader category referred to as anything as a service (XaaS). XaaS is a collective term that refers to the delivery of anything as a service. It refers to the vast number of products and technologies that can be delivered to users as a service over the Internet rather than being provided locally or on-premises. IaaS, PaaS, and SaaS all fall under the broader XaaS category. However, several other products and technologies that are delivered as services are gaining traction in the IT industry. A couple of these other common cloud services models are backup as a service (BaaS) and disaster recovery as a service (DRaaS).

BaaS is a method of offsite data storage where files, folders, disks, or entire computers (including virtual machines) are regularly backed up to cloud services provider infrastructure. The cloud services providers provide a cloud-based data repository, and the cloud services client connects to the repository over the Internet or over a dedicated network. BaaS offers several advantages, including having a copy of data offsite in the event that a natural disaster strikes, and provides the added benefit of offloading backup maintenance and management to the cloud services provider. Many cloud services providers offer BaaS, and software such as VEEAM Cloud Connect facilitates cloud-based repositories.

DRaaS replicates and hosts servers (mostly virtual) on cloud services provider infrastructure to provide failover capabilities in the event of a business disruption due to power failure or natural disaster. DRaaS offers small or large companies disaster recovery (DR) capabilities in an offsite environment. Although it is common for companies to replicate on-premises servers to a cloud environment, it is also possible with DRaaS to replicate in a multi-cloud environment—that is, replicating servers from cloud to cloud. Many cloud services providers provide DRaaS, and software such as VEEAM Cloud Connect and Zerto facilitates a replication environment for the workload to be replicated.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. What are the basic cloud services models identified by NIST? (Choose three.)
   - ❍ **A.** SaaS
   - ❍ **B.** BaaS
   - ❍ **C.** IaaS
   - ❍ **D.** PaaS

2. Which cloud services model allows for the delivery of computing infrastructure where the cloud services consumers can define the operating systems and application?
   - ❍ **A.** SaaS
   - ❍ **B.** BaaS
   - ❍ **C.** IaaS
   - ❍ **D.** PaaS

3. Which cloud services model allows the cloud services consumer to deploy applications on top of a predefined stack defined by the cloud services provider?
   - ❍ **A.** SaaS
   - ❍ **B.** BaaS
   - ❍ **C.** IaaS
   - ❍ **D.** PaaS

# Answers

1. **A**, **C**, and **D** are correct. Software as a service, infrastructure as a service, and platform as a service are the three main cloud services models defined by NIST.
2. **C** is correct. Infrastructure as a service gives the cloud services consumers the ability to provision compute, storage, and network resources, as well as their own operating systems and applications.
3. **D** is correct. Platform as a service allows cloud services consumers to build on top of the cloud services provider infrastructure by using programming languages, libraries, services, and tools.

# Cloud Deployment Models

The cloud delivery or deployment models determine how cloud services consumers use cloud solutions, where the infrastructure is hosted, and who operates the infrastructure. Cloud computing supports four cloud delivery or deployment models:

- **Private cloud:** In a private cloud, the cloud infrastructure is offered for use exclusively by a single organization. The private cloud model allows a single organization to experience cloud computing benefits without sharing underlying resources with other tenants. A private cloud gives a sense of better control and security over an infrastructure. It can be owned, managed, and operated by an organization or a third party over the Internet or over a dedicated link. There are two types of private cloud deployment models:
  - **Internal private cloud:** In an internal private cloud, the infrastructure is on-premises; an internal private cloud is managed by the organization internally.
  - **Hosted private cloud:** In a hosted private cloud, a third-party cloud services provider hosts the private cloud infrastructure. The hosted private cloud is an off-premises approach, meaning it is not physically hosted on the premises of the organization that is using the cloud resources.
- **Public cloud:** This cloud infrastructure is provisioned for use by multiple customers. It may be owned, managed, and operated by a business, academic, or government organization. A cloud services provider's cloud services often run in a remote data center, and the cloud services consumers access it over the Internet or using a dedicated link. The compute, network, or storage resources that make up a public cloud infrastructure are shared among different customers. Still, the customer data remains isolated from other customers; they share the same physical underlying infrastructure but remain isolated due to the multitenant nature of public clouds. Because cloud services providers use a shared infrastructure for tenants, they can realize cost savings, which can be passed on to consumers. Public cloud infrastructure includes IaaS, PaaS, and SaaS.
- **Hybrid cloud:** A hybrid cloud infrastructure combines two or more types of cloud environments. The cloud could be a combination of private, public, and community cloud environments and may also include an on-premises legacy infrastructure. These different cloud infrastructures are unique but are tightly interconnected, almost functioning as one single infrastructure. An organization can use its private cloud for some

workloads and a public cloud for other workloads; it may use the public cloud as a backup for the private cloud or use a public cloud to handle periods of high demands in their private cloud.

- **Community cloud:** A community cloud infrastructure is set up for use by a specific community of users from organizations with shared concerns or common interests, including mission, security requirements, policy, and compliance considerations. A community cloud infrastructure may be hosted on-premises or in a third-party data center. It may be owned, managed, and operated by one or more of the participating organizations in the community or by a third party. A community cloud can provide the privacy and security benefits of a private cloud while providing the pay-as-you-go billing model and cost savings of the public cloud.

Figure 21.5 highlights the main features of the various cloud deployment or delivery models.

**Cloud Deployment or Delivery Models**

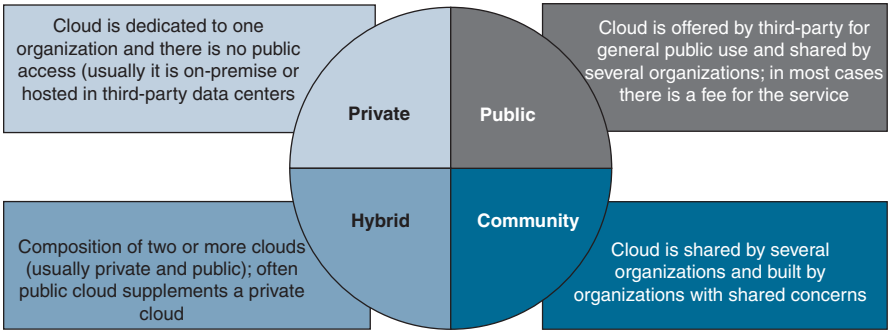



FIGURE 21.5 **Cloud Deployment or Delivery Models**

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which of the following are primary cloud deployment models? (Choose three.)
   - ❍ **A.** Hybrid cloud
   - ❍ **B.** Private cloud
   - ❍ **C.** Public cloud
   - ❍ **D.** Isolated cloud

2. Which cloud deployment model provides the best control and security over an infrastructure?
   - ❍ **A.** Hybrid cloud
   - ❍ **B.** Private cloud
   - ❍ **C.** Community cloud
   - ❍ **D.** Public cloud

3. Which of the following is a feature of a community cloud?
   - ❍ **A.** It is provisioned for use exclusively by a single organization.
   - ❍ **B.** It combines two or more cloud infrastructures.
   - ❍ **C.** It is open for general public usage.
   - ❍ **D.** It is provisioned for use by a specific community of consumers.

# Answers

1. **A**, **B**, and **C** are correct. Hybrid cloud, private cloud, and public cloud are all primary cloud deployment or cloud delivery models.
2. **B** is correct. A private cloud gives a sense of better control and security over an infrastructure.
3. **D** is correct. A community cloud is a cloud infrastructure set up for use by a specific community of users from organizations with shared concerns or common interests.

# On-Premises or Cloud Infrastructure

**ExamAlert**

Before taking the ENCOR exam, make sure you understand the considerations involved in choosing an on-premises or cloud deployment.

Several factors need to be considered to determine if a cloud infrastructure would be a better fit for an organization than an on-premises infrastructure. The following lists highlight some of the primary considerations for deploying each model.

On-premises considerations:

- Provides a sense of being more secure because applications and data are on an organization's own server, in its own storage, and behind its own firewalls.
- The organization may be operating in a highly regulated environment with requirements for applications and data to be hosted on-premises.
- The organization may incur higher costs because an on-premises infrastructure requires hardware, software licenses, and dedicated IT staff to manage and troubleshoot the infrastructure.
- Complete customization of the infrastructure is possible because the organization is totally in control of all the pieces of its on-premises infrastructure.

Cloud considerations:

- Cloud services consumers can use a pay-as-you-go billing model, eliminating capital expenses for infrastructure setup.
- Cloud infrastructure can quickly scale up and down to a nearly unlimited capacity, depending on the consumer usage growth.
- Cloud infrastructure provides near-instant provisioning, reducing installation and configuration time and bringing applications to cloud services consumers' applications.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. True or false: Higher costs may be incurred when deploying infrastructure completely on-premises.
   - ❍ **A.** True
   - ❍ **B.** False

2. True or false: Cloud-based deployments provide a pay-as-you-go billing model.
   - ❍ **A.** True
   - ❍ **B.** False

## Answers

1. **A** is correct. An organization may incur higher costs when deploying infrastructure completely on-premises as on-premises infrastructure requires hardware, software licenses, and dedicated IT staff to manage and troubleshoot the infrastructure.
2. **A** is correct. Cloud services consumers can use a pay-as-you-go billing model, which eliminates capital expenses for infrastructure setup.

# Review Questions

1. Which cloud services model gives a cloud services consumer the ability to use applications hosted on the cloud services provider's infrastructure?
   - ❍ **A.** SaaS
   - ❍ **B.** BaaS
   - ❍ **C.** IaaS
   - ❍ **D.** PaaS

2. What characteristics does NIST associate with cloud services? (Choose three.)
   - ❍ **A.** On-demand self-service
   - ❍ **B.** Broad network access
   - ❍ **C.** Resource reservation
   - ❍ **D.** Resource pooling

3. Which type of private cloud can be hosted off-premises?
   - ❍ **A.** Community cloud
   - ❍ **B.** Hybrid cloud
   - ❍ **C.** Hosted private cloud
   - ❍ **D.** Internal private cloud

4. Which of the following is a consideration for keeping an infrastructure on-premises?
   - ❍ **A.** A sense of nearly unlimited capacity
   - ❍ **B.** Support for a pay-as-you-go model
   - ❍ **C.** Faster time to market
   - ❍ **D.** Operating in a highly regulated environment

# Answers to Review Questions

1. **A** is correct. With software as a service, the cloud services provider makes available its application for use by the cloud services consumer.
2. **A**, **B**, and **D** are correct. On-demand self-service, broad network access, and resource pooling are all essential characteristics that NIST associates with cloud services.
3. **C** is correct. With a hosted private cloud, a third-party cloud services provider hosts the private cloud infrastructure.
4. **D** is correct. When organizations operate in a highly regulated environment, they may have requirements for applications and data to be hosted on-premises.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*
- *NIST Special Publication (SP) 800-145, The NIST Definition of Cloud Computing: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers SD-WAN.

CHAPTER 22

# SD-WAN

**This chapter covers the following official ENCOR 350-401 exam objectives:**

- 1.4 Explain the working principles of the Cisco SD-WAN solution
- 1.4.a SD-WAN control and data planes elements
- 1.4.b Traditional WAN and SD-WAN solutions

Cisco Software-Defined Wide Area Network (SD-WAN) is one of two technologies that are part of Cisco's software-defined networking (SDN) solution used in enterprise networks. The other, SD-Access, is covered in Chapter 23, "SD-Access." SDN is a centralized approach to managing networks that abstracts the underlying network infrastructure from its applications. Cisco SD-WAN is a WAN architecture overlay that enables digital and cloud transformation for enterprises that integrate routing, security, centralized policy, and orchestration into large-scale networks.

This chapter provides an overview of Cisco SD-WAN and describes its architecture and components. The first section looks at the need for SD-WAN and the benefits SD-WAN provides over traditional WANs. The next section looks at the architecture and components of an SD-WAN solution. It covers the various planes of operation, including the control plane, data plane, orchestration plane, and management plane. This chapter also looks at the various SD-WAN components that operate at these planes.

**This chapter covers the following technology topics:**

- SD-WAN Overview
- SD-WAN Architecture and Components

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. What is the name of the Cisco SD-WAN feature that automates connectivity to workloads in the public cloud when accessed from a data center or branch?
2. What two SD-WAN router options are part of the SD-WAN solution?

### Answers

1. Cisco Cloud OnRamp for IaaS
2. vEdge and cEdge

# SD-WAN Overview

As the enterprise landscape continuously evolves, the need and demand for mobile and Internet of Things (IoT) device traffic, software as a service (SaaS) applications, and cloud adoption grow. In addition, the security needs of organizations are increasing, and applications are requiring prioritization and optimization.

With the rise of cloud deployment models like infrastructure as a service (IaaS), traffic patterns are changing. The use of cloud applications and services like Microsoft Office 365, Amazon Web Services (AWS), and Microsoft Azure changes traffic patterns. The majority of enterprise traffic now flows to public clouds and the Internet. These changes are creating new requirements for security, application performance, cloud connectivity, WAN management, and operations that traditional WAN solutions, like MPLS, were not designed to address.

The Cisco SD-WAN solution addresses these challenges. Using the abstraction that SDN provides, SD-WAN separates the data plane and control plane to centralize the intelligence of the network. It allows for network automation, operations simplification, centralized provisioning, monitoring, and troubleshooting. SD-WAN can support multitenancy and is cloud delivered, highly automatable, scalable, and application aware. Cisco's approach to SDN with SD-WAN is based on technology from its acquisition of Viptela.

Figure 22.1 shows the flexibility of Cisco SD-WAN delivery.

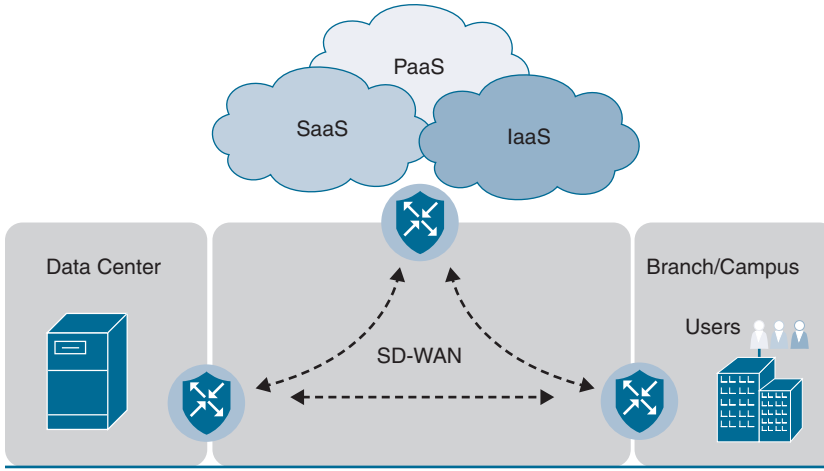



FIGURE 22.1 **Cisco SD-WAN Delivery**

**ExamAlert**

For the ENCOR exam, make sure you understand the need for SD-WAN and the advantages it provides over legacy WAN deployments.

# The Need for SD-WAN

Legacy WAN architectures face challenges due to the evolving landscape. A typical legacy WAN architecture may have multiple MPLS transports, or an MPLS transport paired with an Internet link, or LTE used in an active/backup fashion. Often, Internet or SaaS traffic is backhauled to a central data center/headquarters for Internet access. Issues with these architectures include insufficient bandwidth and high bandwidth costs, application downtime, poor SaaS performance, complex operations, complex workflows for cloud connectivity, long deployment times and policy changes, limited application visibility, and difficulty securing the network.

Cisco SD-WAN represents a shift from the legacy hardware-based WAN design to a more secure, software-driven virtual IP fabric overlay that runs over standard network transport services. The Cisco SD-WAN solution creates an overlay network that builds secure, unified connectivity over any transport network (the underlay). The underlay network, which is the physical infrastructure for the WAN, can be the public Internet, MPLS, Metro Ethernet, LTE/4G/5G, or some other service. The underlay network facilitates the creation of overlay networks and is responsible for delivering packets across networks.

The following are some common use cases for the deployment of Cisco SD-WAN over traditional WANs:

- **Secure automated WAN:** SD-WAN provides secure connectivity between remote offices, data centers, and public/private clouds over a transport-independent network.
- **Application performance optimization:** SD-WAN improves the application experience for users at remote offices.
- **Secure Direct Internet Access (DIA):** SD-WAN locally offloads Internet traffic at the remote office.
- **Multicloud:** SD-WAN connects remote offices with cloud (SaaS and IaaS) applications over an optimal path and through regional co-location/exchange points where security services can be applied.

The following sections look more closely at these use cases.

# Secure Automated WAN

SD-WAN provides secure connectivity between remote offices, data centers, and public/private clouds. It allows you to streamline the automation of a WAN with these features:

- **Automated zero-touch provisioning:** It is possible to automate edge router provisioning. An edge router discovers its controller automatically and, once authenticated, automatically downloads the configuration before establishing IPsec tunnels with the rest of the network.
- **Bandwidth augmentation:** SD-WAN makes it possible to increase WAN bandwidth by leveraging all available WAN transports and routing capabilities to distribute traffic across available paths in an active/active fashion. Traffic can be offloaded from more expensive circuits (such as MPLS transports) to broadband circuits, which can achieve the same availability.
- **VPN segmentation:** Traffic isolation is critical in the SD-WAN security strategy. As traffic enters a router, it is assigned to a VPN. This allows for both user traffic isolation and routing table isolation. Traffic isolation ensures that a user in one VPN cannot transmit data to another VPN unless such transmission is explicitly configured.

- **Centralized management:** The vManage interface offers centralized configuration, accounting, performance, and security management in a single pane of glass. It provides operational simplicity and streamlines deployment through the use of ubiquitous policies and templates. It therefore reduces change control and deployment times.

## Application Performance Optimization

The following SD-WAN capabilities help address application performance optimization to improve the application experience for users:

- **Application-aware routing:** SD-WAN enables you to customize service level agreement (SLA) policies for traffic and measure real-time performance with Bidirectional Forwarding Detection (BFD) probes. During periods of performance degradation, traffic can be directed to other paths that support the SLA for that application if SLAs on the original path are exceeded.
- **Quality of service (QoS):** SD-WAN allows for the classification, scheduling, queueing, shaping, and policing of traffic on the WAN router interfaces. Its QoS features are designed to minimize the delay, jitter, and packet loss of critical application flows.
- **Forward error correction (FEC) and packet duplication:** SD-WAN allows for packet loss mitigation. With FEC, the transmitting WAN edge inserts a parity packet for every four data packets, and the receiving WAN edge can reconstruct a lost packet based on the parity value. With packet duplication, the transmitting WAN edge replicates all packets for selected critical applications over two tunnels at a time, and the receiving WAN edge reconstructs critical application flows and discards the duplicate packets.
- **TCP optimization and session persistence:** SD-WAN addresses high latency and poor throughput for long-haul or high-latency satellite links. With TCP optimization, a WAN edge router acts as a TCP proxy between a client and a server. With session persistence, instead of a new connection for every single TCP request and response pair, a single TCP connection is used to send and receive multiple requests and responses.

# Secure Direct Internet Access (DIA)

Backhauling traffic to a central site causes increased bandwidth utilization for security and network devices and links at the central site. This also increases latency, which impacts application performance. DIA helps solve these issues by allowing Internet-bound traffic from a VPN to exit the remote site locally.

However, DIA can pose some security challenges as remote site traffic needs security against Internet threats. With Cisco SD-WAN, you can address this issue by leveraging the embedded security features on IOS XE SD-WAN devices or using Secure Internet Gateway (SIG) or Cisco Umbrella. IOS XE SD-WAN provides a number of security features, including an application-aware firewall, intrusion detection systems (IDSs)/intrusion prevention systems (IPSs), DNS/web layer security, URL filtering, SSL proxies, and Advanced Malware Protection (AMP).

# Multicloud

Cisco Multicloud connects IaaS or SaaS cloud applications to remote sites over optimal paths and through regional co-location/exchange points where security services can be applied centrally. Cisco SD-WAN Multicloud supports the following use cases:

- **Infrastructure as a service (IaaS):** IaaS delivers network, compute, and storage resources to remote users as a service over the Internet. These services are often available in a public cloud (such as AWS or Azure). Traditionally, for a branch to reach IaaS resources, there would not be any direct access to these public clouds; rather, access would be through a data center or co-location site. If there is a dependency on MPLS to reach IaaS resources, oftentimes there is not consistent segmentation and there are not QoS policies for traffic from the branch to the public cloud. Cisco Cloud OnRamp can help alleviate some of these issues.

  Cisco Cloud OnRamp for IaaS is a feature that automates connectivity to workloads in the public cloud when accessed from a data center or branch. WAN edge router instances are automatically deployed in the public cloud and become part of the SD-WAN network overlay. OnRamp establishes data plane connectivity back to the data center or branch routers, eliminating the need for traffic from SD-WAN sites to traverse the data center. This improves the performance of applications hosted in the cloud as well as their availability. High availability and path redundancy are achieved because a pair of virtual routers is deployed in AWS VPC or Azure VNet.

- **Software as a service (SaaS):** SaaS is a cloud deployment model in which a cloud provider hosts an application and provides access to the application to end users over the Internet. When accessing SaaS applications (for example, Office 365, Box) through a centralized data center, access to the applications may be susceptible to increased latency, and users may have an unpredictable experience. Cisco SD-WAN gives you in-depth visibility into SaaS applications when using DIA for access from a remote site.

  Using Cloud OnRamp for SaaS, you can enhance the end-user experience by choosing the network path for a particular SaaS application. In the case of a change in the network due to impairment, loss, or delay, the SaaS traffic can dynamically and intelligently move over to the next optimal path.

- **Regional Multicloud access:** When regulatory compliance or a company's security policy does not allow DIA at a remote site, you may find that using Cloud OnRamp for Colocation is feasible. Cloud OnRamp for Colocation allows for a hybrid approach, with colocation centers at strategic points in the network to consolidate network and security stacks and minimize latency.

  Colocation centers are public data centers where organizations can rent equipment space and connect to a variety of network and cloud service providers. Colocation centers are purposely selected for close proximity to end users and for high-speed access to public and private cloud resources. They are more cost-effective than a dedicated private data center.

Figure 22.2 shows Cisco SD-WAN Multicloud options.

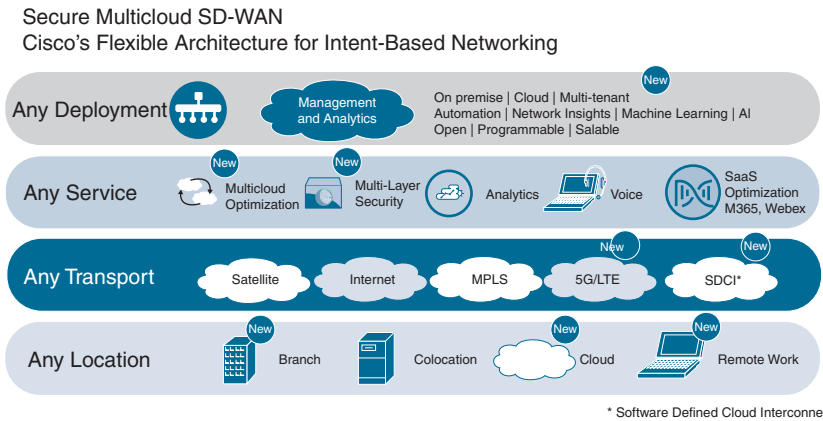



FIGURE 22.2 **Cisco SD-WAN Multicloud Options**

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which SD-WAN feature enables you to choose what network path a particular cloud-based application traffic goes over?
   - ❍ **A.** Cisco Cloud OnRamp for IaaS
   - ❍ **B.** Cisco Cloud OnRamp for SaaS
   - ❍ **C.** Cisco Cloud OnRamp for Colocation
   - ❍ **D.** Cisco Cloud OnRamp for PaaS

2. Which of the following components is the single pane of glass for the SD-WAN solution?
   - ❍ **A.** Cisco ISE
   - ❍ **B.** Cisco DNA Center
   - ❍ **C.** vSmart
   - ❍ **D.** vManage

## Answers

1. **B** is correct. Using Cisco Cloud OnRamp for SaaS, you can choose the network path a particular SaaS application uses for traffic.
2. **D** is correct. vManage offers centralized configuration, accounting, performance, and security management in a single pane of glass.

# SD-WAN Architecture Components

This section looks at the architecture components of a Cisco SD-WAN solution. The Cisco SD-WAN solution is based on the same routing principles that have been in use for years. However, Cisco SD-WAN separates the data plane from the control plane and virtualizes much of the routing that would otherwise require dedicated hardware. The separation between the control and data planes enables the Cisco SD-WAN solution to run over any transport circuits. The control plane manages the rules for routing traffic through the overlay network, while the data plane moves data packets across the network devices.

**ExamAlert**

For the ENCOR exam, make sure you understand the SD-WAN planes of operation and the components within them.

These planes of operation are involved in an SD-WAN solution:

- **Control plane:** The control plane builds and maintains the network topology and makes forwarding decisions about where traffic flows.
- **Data plane:** The data plane is responsible for forwarding packets based on decisions made from the control plane.
- **Orchestration plane:** The orchestration plane assists in the automatic onboarding of the SD-WAN routers into the SD-WAN overlay fabric.
- **Management plane:** The management plane is responsible for central configuration and monitoring of an SD-WAN deployment.

The main components of a SD-WAN solution include the vSmart controller (control plane), the WAN edge router (data plane), the vBond orchestrator (orchestration plane), and the network management system (management plane). The following sections look at these components in more detail.

## vSmart Controllers

A vSmart controller is a software-based component that is responsible for the centralized control plane of an SD-WAN network. It maintains a secure connection to each WAN edge router and distributes routes and policy information via Overlay Management Protocol (OMP), thereby acting as a route reflector. It also orchestrates secure data plane connectivity between the WAN edge routers by reflecting crypto key information originating from WAN edge routers, allowing for a scalable, IKE-less architecture.

You can think of the vSmart controllers as the brains of an SD-WAN solution. They have pre-installed credentials that allow them to authenticate the SD-WAN routers as they come online. Using these credentials guarantees that only authenticated devices are allowed access to the SD-WAN fabric. With successful authentication, each vSmart controller establishes a permanent DTLS tunnel to each SD-WAN router in the SD-WAN fabric. These tunnels are then used to establish OMP neighborships with each SD-WAN router. (We take a closer look at OMP toward the end of this chapter.)

A vSmart controller processes the OMP routes learned from the SD-WAN routers to decide on the network topology and calculate the best routes to various network destinations. It then advertises the reachability information learned from these routes to all SD-WAN routers in the fabric. A vSmart controller behaves similarly to a BGP route reflector. It receives routes from WAN edge routers and then processes and applies any policy to them. Finally, it advertises the routes to other WAN edge routers in the overlay network.

# WAN Edge Routers

WAN edge routers are available as either hardware appliances or software-based routers. They can sit at a physical site or in the cloud, and they provide secure data plane connectivity among the sites over one or more WAN transports. When deployed as a virtual appliance, a WAN edge router can be hosted in a private cloud environment such as VMware ESXi or KVM or in a public cloud such as AWS or Azure. A WAN edge router is responsible for traffic forwarding, security, encryption, QoS, and routing protocols such as BGP and OSPF.

Two different SD-WAN router options are available for Cisco SD-WAN:

- **vEdge:** This is the original Viptela platform running Viptela software.
- **cEdge:** This is the Viptela software integrated with Cisco IOS XE. Supported platforms are CSR, ISR, ASR1K, and ENCS platforms, as well as the cloud-based CSRv and ISRv.

The IOS XE image for the SD-WAN platform is slightly different from the standard IOS XE image. Only a selected set of IOS XE features that makes sense for an SD-WAN environment are available with the SD-WAN IOS XE image.

WAN edge routers use OSPF and BGP routing protocols to learn reachability information. WAN edge routers have a mature feature set of routing implementations that accommodates simple, moderate, and complex routed environments. For redundancy, you can have the WAN edge routers implement the first-hop redundancy protocol (FHRP) Virtual Router Redundancy Protocol (VRRP), which can operate on a per-VLAN basis.

WAN edge routers are responsible for encrypting and decrypting application traffic between the sites. WAN edge routers establish a control plane relationship with the vSmart controller to exchange the information required to establish the fabric. WAN edge routers also export performance statistics, alerts, and events to the vManage system for management purposes.

## vBond Orchestrators

A vBond orchestrator is a software-based component that performs the initial authentication of WAN edge devices and orchestrates vSmart, vManage, and WAN edge connectivity. It also enables communication between devices that sit behind Network Address Translation (NAT).

The orchestration plane handles the following tasks:

- **Joining the WAN edge to the overlay:** For the WAN edge router to join the overlay network, it needs to establish a secure connection to the vManage to receive a configuration file. It also needs to establish a secure connection with the vSmart controller to participate in the overlay network. The discovery of vManage and vSmart happens automatically and is accomplished by first establishing a secure connection to the vBond orchestrator.
- **Onboarding the WAN edge router:** You can manually configure or automate the process of getting WAN edge routers up and running. When using the manual method, you establish a console connection to run a few lines of configuration commands. In automated provisioning mode, such as with zero-touch provision or plug-and-play, you can plug a WAN edge router into the network and power it on, and it is provisioned automatically. With IOS XE SD-WAN routers only, you have the option of using the bootstrap method, where there is a configuration loaded via bootflash or a USB key to get the device onto the SD-WAN network. Onboarding virtual cloud routers involves configuring a one-time password to get temporarily authenticated before device certificates can be permanently obtained through vManage.

## vManage

vManage is the centralized network management system for SD-WAN. It is software based and provides a single-pane-of-glass view through the GUI interface. Using the GUI interface, you can easily monitor, configure, and maintain all Cisco SD-WAN devices and their connected links in the underlay and overlay networks.

Some key features of Cisco vManage include centralized provisioning, centralized policies, device configuration templates, and the ability to monitor an entire SD-WAN solution. You can also use vManage to perform centralized software upgrades on all fabric elements, including WAN edge, vBond, vSmart, and vManage. You can also export performance statistics to an external system or to the Cisco vAnalytics tool for further examination. vAnalytics is an optional analytics and assurance service that provides visibility into application and infrastructure across the WAN, forecasting, and what-if analysis.

Cisco SD-WAN software provides the REST API programmatic interface for controlling, configuring, and monitoring Cisco SD-WAN devices. The vManage network management system (NMS) web server uses HTTP or HTTPS as the communications protocol. The API plays a critical role for clients to consume the features that are provided by vManage. These are some of the common use cases for the vManage API:

- Monitoring device status
- Configuring a device
- Querying and aggregating device statistics

Figure 22.3 shows the Cisco vManage main dashboard.

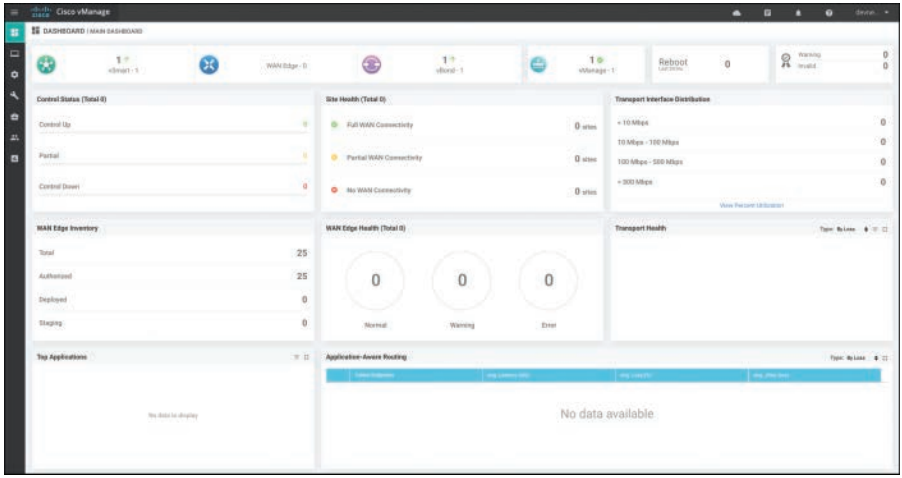



FIGURE 22.3 **The Cisco vManage Main Dashboard**

# SD-WAN Considerations

To wrap up this chapter, let us take a look at some elements that you need to consider for a SD-WAN deployment:

- **IP addresses:** There are two main types of IP addresses:
  - **Private IP address:** On WAN edge routers, a private IP address is an IP address assigned to the interface of the SD-WAN device. This is the pre-NAT address, and despite the name, it can be a public address (that is, publicly routable) or a private address.
  - **Public IP address:** This is the post-NAT address detected by the vBond orchestrator. This address can be either a public address (that is, publicly routable) or a private address.
- **NAT transversal:** An SD-WAN router may be unknowingly sitting behind a NAT device. It is important to know what IP address/port to connect to from outside the network in order to establish connections in the SD-WAN network. vBond acts as a Session Traversal Utilities for NAT (STUN) server, allowing other controllers and SD-WAN routers to discover their own mapped/translated IP addresses and port numbers.
- **Transport locators (TLOCs):** A TLOC is an attachment point where a WAN edge router connects to the WAN transport network. It is uniquely identified and represented by a three-tuple, consisting of the system IP address, link color, and encapsulation (Generic Routing Encapsulation or IPsec).
- **Overlay Management Protocol (OMP):** The OMP routing protocol has a similar structure to BGP and manages the SD-WAN overlay network. Its protocol runs between vSmart controllers and between vSmart controllers and WAN edge routers, where control plane information—such as route prefixes, next-hop routes, crypto keys, and policy information—is exchanged over a secure DTLS or TLS connection.
- **Virtual private networks (VPNs):** In the SD-WAN overlay, VPNs provide segmentation similar to virtual routing and forwarding (VRF) instances. VPNs are isolated from one other, and each has its own forwarding table. An interface or subinterface is explicitly configured under a single VPN and cannot be part of more than one VPN. Labels are used in OMP route attributes and packet encapsulation, which identifies which VPN a packet belongs to.

For the vBond orchestrator, only two VPNs are functional and should be used:

- **VPN 0:** This is the transport VPN. It contains the interfaces that connect to the WAN transports.
- **VPN 512:** This is the management VPN. It carries the out-of-band management traffic to and from the Cisco SD-WAN devices.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. In an SD-WAN solution, which plane of operation is responsible for the automatic onboarding of SD-WAN routers?
   - ❍ **A.** Control plane
   - ❍ **B.** Data plane
   - ❍ **C.** Orchestration plane
   - ❍ **D.** Management plane

2. True or false: WAN edge routers can only be deployed as physical appliances.
   - ❍ **A.** True
   - ❍ **B.** False

## Answers

1. **C** is correct. The orchestration plane assists in the automatic onboarding of the SD-WAN routers into the SD-WAN overlay fabric.
2. **B** is correct. A WAN edge router is available as either a hardware appliance or a software-based router. It can be hosted in a private cloud environment such as VMware ESXi or KVM, or in a public cloud, such as AWS or Azure.

# Review Questions

1. Which SD-WAN feature automatically deploys WAN edge router instances in the public cloud to become part of the SD-WAN network overlay?
   - ❍ **A.** Cisco Cloud OnRamp for IaaS
   - ❍ **B.** Cisco Cloud OnRamp for SaaS
   - ❍ **C.** Cisco Cloud OnRamp for Colocation
   - ❍ **D.** Cisco Cloud OnRamp for PaaS

2. In an SD-WAN solution, which plane of operation is responsible for central configuration and monitoring of the SD-WAN deployment?
   - ❍ **A.** Control plane
   - ❍ **B.** Data plane
   - ❍ **C.** Orchestration plane
   - ❍ **D.** Management plane

3. Which SD-WAN component acts similarly to a BGP route reflector?
   - ❍ **A.** vSmart controller
   - ❍ **B.** WAN edge router
   - ❍ **C.** vBond orchestrator
   - ❍ **D.** vManage

# Answers to Review Questions

1. **A** is correct. Cisco Cloud OnRamp for IaaS allows for the automatic deployment of WAN edge router instances in the public cloud that become part of the SD-WAN network overlay.
2. **D** is correct. The management plane is responsible for central configuration and monitoring of an SD-WAN deployment.
3. **A** is correct. A vSmart controller behaves similarly to a BGP route reflector. It receives routes from WAN edge routers, processes and applies any policy to them, and then advertises the routes to other WAN edge routers in the overlay network.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*
- *Cisco Software-Defined Wide Area Networks: Designing, Deploying and Securing Your Next Generation WAN with Cisco SD-WAN*
- *CCNP Enterprise Design ENSLD 300-420 Official Cert Guide: Designing Cisco Enterprise Networks*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers SD-Access.

CHAPTER 23

# SD-Access

**This chapter covers the following official ENCOR 350-401 exam objectives:**

- 1.5 Explain the working principles of the Cisco SD-Access solution
- 1.5.a SD-Access control and data planes elements
- 1.5.b Traditional campus interoperating with SD-Access

This chapter covers the benefits of SD-Access over traditional campus networks. It looks at the layers of an SD-Access architecture and covers the different planes of operation for an SD-Access fabric. This chapter examines the Locator/ID Separation Protocol (LISP) control plane, the Virtual Extensible LAN (VXLAN) data plane, the Cisco TrustSec policy plane, and Cisco DNA Center management plane. Although LISP and VXLAN are covered exclusively in Chapter 28, "Extending the Network Virtually," this chapter covers their use in terms of SD-Access. Likewise, TrustSec is covered in more depth in Chapter 10, "Network Security Design." It may be worthwhile to review the related sections of those chapters before reading through this one. Finally, this chapter looks at the fabric roles and components of the SD-Access architecture.

**This chapter covers the following technology topics:**

- SD-Access Overview
- SD-Access Architecture
- SD-Access Fabric Roles and Components

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. What collective grouping of devices is responsible for transporting packets between network devices for the SD-Access fabric overlay?
2. What SD-Access architecture component is part of the Layer 3 network used for interconnections between border nodes and edge nodes?

### Answers

1. Underlay
2. Intermediate node

# SD-Access Overview

Cisco Software-Defined Access (SD-Access) is one of two technologies that are part of Cisco's software-defined networking (SDN) solution used in the enterprise. The other, SD-WAN, is covered in Chapter 22, "SD-WAN." SD-Access is the evolution of the typical enterprise campus design to a modern design that truly represents the changing needs of the network and the intent of the organization. The SD-Access solution runs on Cisco Digital Network Architecture (DNA) Center hardware to automate both wired and wireless networks.

The fabric technology, a crucial part of SD-Access, provides campus networks with programmable overlay networks and simplified network virtualization. This facilitates the creation of one or more logical or overlay networks that meet the design intent on a physical or underlay network. The fabric technology also allows for tighter security, providing software-defined segmentation and enforcement of policies based on user identity and group membership.

In SD-Access, software-defined segmentation integrates Cisco TrustSec technology, allowing for microsegmentation for groups within a virtual network using security group tags (SGTs). Using Cisco DNA Center with a single dashboard, you can automate network creation with integrated security and segmentation to reduce risks, operational complexity, and expenses. Cisco Identity Services Engine (ISE) integrates with SD-Access, enabling the dynamic mapping of users to scalable groups and simplifying end-to-end policy enforcement. Finally, by using the DNA Center, you can gather intelligence related to

network performance, network insights, and telemetry through the assurance and analytics capabilities.

**ExamAlert**

For the ENCOR exam, make sure you understand why there was a need for SD-Access.

Over the years, with digitization, many applications evolved from simply supporting business processes to becoming, in some cases, primary revenue sources and competitive differentiators. It is now necessary to scale the capacity of a network to meet the demands and growth of applications. High administrative overhead is involved in scaling traditional networks. In addition, it is now necessary to manage various levels of access of users and their devices to a variety of applications. Wired and wireless networks now need to be able to scale to support the changing demands of end users. The traditional campus network cannot scale so quickly, is prone to errors in configuration, and does not provide the flexibility to meet these new demands.

SD-Access represents a significant change in how to design, provision, operate, and troubleshoot the enterprise campus network environment. It reduces dependency on manual CLI configurations that can lead to risk and network downtime due to misconfiguration. Also, traditional network management is hardware-centric, is tedious, and does not scale well in this era of digitization. The SD-Access architecture offers open and standards-based APIs for management, as well as a single pane of glass (DNA Center) for managing the enterprise campus network environment.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. SD-Access can be considered an evolution of which of the following network environments?
   - ❍ **A.** Data center
   - ❍ **B.** Enterprise campus
   - ❍ **C.** WAN
   - ❍ **D.** Private cloud

2. Which component integrates with the SD-Access environment to enable the dynamic mapping of users to scalable groups for end-to-end policy enforcement?
   - ❍ **A.** Cisco ISE
   - ❍ **B.** Cisco DNA Center
   - ❍ **C.** Cisco TrustSec
   - ❍ **D.** NMS

# Answers

1. **B** is correct. SD-Access is the evolution from the typical enterprise campus design to a modern design that truly represents the changing needs of the network and the intent of the organization.
2. **A** is correct. Cisco ISE integrates with SD-Access, enables the dynamic mapping of users to scalable groups, and simplifies end-to-end policy enforcement.

# SD-Access Architecture

At a high level, the SD-Access architecture consists of both hardware and software components. The SD-Access fabric architecture comprises four different layers. Throughout this chapter, we touch on these layers and their subcomponents:

- **Physical layer:** SD-Access needs to run on physical hardware. All of the Cisco hardware that participates in SD-Access needs to support all of the hardware application-specific integrated circuit (ASIC) and field-programmable gate array (FPGA) requirements. The SD-Access fabric includes these physical layer devices:
  - **Cisco switches:** Switches provide wired access to the fabric. Multiple types of Cisco Catalyst switches are supported, such as Catalyst 9500 and 3850.
  - **Cisco routers:** Routers provide WAN and branch access to the fabric. Multiple types of Cisco routing platforms are supported, such as ASR 1000, ISR 4400, and CSRv.
  - **Cisco wireless:** Cisco WLCs and APs provide wireless access to the fabric. Multiple WLC and AP platforms are supported, such as C9800 WLC and C9100 APs.
  - **Cisco controller appliances:** Cisco DNA Center and Cisco ISE are the two controller appliances required for an SD-Access deployment.
- **Network layer:** The network layer is made up of the underlay network and the overlay network. These layers work together in moving packets between network devices that participate in SD-Access:
  - **Underlay:** This underlying physical layer is responsible for transporting packets between network devices for the SD-Access fabric overlay. The underlay network can be manually configured or can be automatically configured using the Cisco DNA Center LAN Automation feature.
  - **Overlay:** This virtual (tunneled) network virtually interconnects all of the network devices, forming a fabric of interconnected nodes of the SD-Access infrastructure. Overlays can be Layer 2 overlays or Layer 3 overlays:
    - **Layer 2 overlays:** A Layer 2 overlay is used for carrying a single subnet over the Layer 3 underlay network. It emulates a LAN segment that can transport IP and non-IP frames.

- **Layer 3 overlays:** A Layer 3 overlay provides network abstraction and allows for multiple IP networks as part of each virtual network. Overlapping IP address space is permitted with Layer 3 overlays, provided that network virtualization is preserved outside the fabric through the use of technologies such as VRF-Lite.

- **Controller layer:** The controller layer is provided by the Cisco DNA Center and Cisco ISE and consists of a number of components:
  - **Cisco Network Control Platform (NCP):** NCP provides all the underlay and fabric automation and orchestration services for the physical and network layers. NCP configures and manages network devices using NETCONF/YANG, Simple Network Management Protocol (SNMP), SSH/Telnet, and so on.
  - **Cisco Network Data Platform (NDP):** NDP is the data collection and analytics subsystem of DNA Center. It analyzes and correlates various network events through multiple sources, such as NetFlow and Switched Port Analyzer (SPAN). The information received is used for providing contextual information to NCP and ISE.
  - **Cisco Identity Services Engine (ISE):** ISE provides network access control (NAC) and identity services for dynamic endpoint-to-group mapping and policy definition in various ways, such as using 802.1x, MAC Authentication Bypass (MAB), and Web Authentication (WebAuth).
- **Management layer:** The management layer is the Cisco DNA Center user interface, where the information from other layers is presented in a centralized management dashboard. The management layer presents a network administrator with a simple set of GUI tools and workflows to manage the entire DNA network. DNA Center hides all the complexities and dependencies of the other layers. Thanks to DNA Center, a complete understanding of all of the other layers' technologies (such as LISP, VXLAN, TrustSec, NCP, NDP, and ISE) is not required to deploy the SD-Access fabric. DNA Center also frees you from needing to know how to configure each network device and its features manually.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which component of an SD-Access architecture provides a centralized dashboard that hides the complexities and dependencies of configuring the other layers?

   - ❍ **A.** Underlay network
   - ❍ **B.** Overlay network
   - ❍ **C.** Cisco DNA Center
   - ❍ **D.** Cisco ISE

2. Which component of the controller layer in an SD-Access architecture provides the underlay and fabric automation and orchestration services for the physical and network layers?

   - ❍ **A.** NCP
   - ❍ **B.** NDP
   - ❍ **C.** Cisco ISE
   - ❍ **D.** LISP

## Answers

1. **C** is correct. Cisco DNA Center provides a centralized management dashboard that hides all the complexities and dependencies of the other layers.
2. **A** is correct. NCP provides all the underlay and fabric automation and orchestration services for the physical and network layers.

# SD-Access Operational Planes

**ExamAlert**

For the ENCOR exam, make sure you understand the various planes of operation and components of SD-Access.

The Cisco SD-Access solution comprises four technologies, each of which performs distinct activities in different planes of operation:

- **Control plane:** This plane involves the messaging and communication protocol between infrastructure devices in the fabric.
- **Data plane:** This plane involves the encapsulation method used for the moving of data packets.
- **Policy plane:** This plane involves the security and segmentation of network devices.
- **Management plane:** This plane involves the orchestration, assurance, visibility, and management of the SD-Access solution.

In SD-Access, the control plane is based on Locator/ID Separation Protocol (LISP). The data plane is based on Virtual Extensible LAN (VXLAN). The policy plane is based on Cisco TrustSec, and Cisco DNA Center enables the management plane. As mentioned earlier in this chapter, some of these technologies are covered in more detail in other chapters. Here we look at how an SD-Access solution uses these technologies:

- **LISP:** LISP allows the separation of identity and location through a mapping relationship of two namespaces: an endpoint identifier (EID) in relation to its routing locator (RLOC). This relationship is known as an EID-to-RLOC mapping. In an SD-Access architecture, rather than making a traditional routing-based decision, the fabric devices query the control plane node to determine the RLOC associated with the destination address (EID-to-RLOC mapping) and then use that RLOC information as the traffic destination. Traffic is sent to the default fabric border node in the event of a failure to resolve the destination RLOC.
- **VXLAN:** VXLAN is a MAC-in-IP encapsulation technique that provides a way to carry lower-layer data across the Layer 3 infrastructure. VXLAN preserves the original Ethernet header from the original frame sent from

an endpoint. This facilitates the creation of an overlay at Layer 2 and Layer 3. SD-Access places additional information in the fabric VXLAN header, including forwarding attributes used to make policy decisions, by identifying each overlay network using a VXLAN network identifier (VNI).

- **TrustSec:** Cisco TrustSec technology is used to assign an SGT value to a packet at its ingress point into a network. Cisco TrustSec tags are assigned to groups of users or devices. Network policies, such as QoS and ACLs, are then applied throughout the SD-Access fabric using the SGT tag instead of using network addresses, such as MAC or IPv4 addresses. This allows for the creation of network policies such as security, QoS, and network segmentation based only on SGT and not on the network address.
- **DNA Center:** Cisco DNA Center enables automation of device deployments and configurations in a network to provide speed and consistency. Thanks to DNA Center's automation capabilities, the control plane, data plane, and policy plane for the fabric devices can be easily, seamlessly, and consistently deployed.

Figure 23.1 shows a snapshot of Cisco DNA Center.

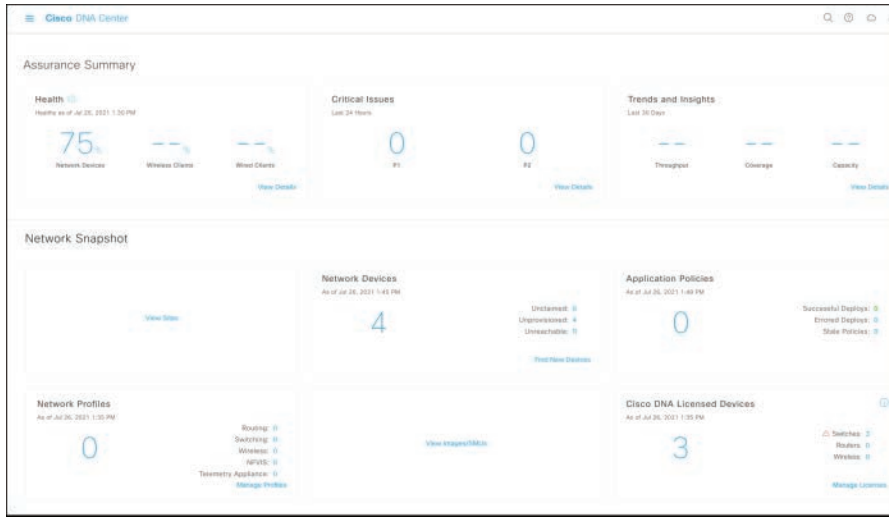



FIGURE 23.1 **Cisco DNA Center**

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. VXLAN functions in which operation plane in an SD-Access architecture?
   - ❍ **A.** Control plane
   - ❍ **B.** Data plane
   - ❍ **C.** Policy plane
   - ❍ **D.** Management plane

2. The assignment of an SGT value as packets enter a network is associated with which operational plane of the SD-Access architecture?
   - ❍ **A.** Control plane
   - ❍ **B.** Data plane
   - ❍ **C.** Policy plane
   - ❍ **D.** Management plane

## Answers

1. **B** is correct. VXLAN is a MAC-in-IP encapsulation technique that provides a way to carry lower-layer data across the Layer 3 infrastructure and operates at the data plane.
2. **C** is correct. At the policy plane, Cisco TrustSec technology assigns an SGT value to the packet at its ingress point into a network.

# SD-Access Fabric Roles and Components

**ExamAlert**

For the ENCOR exam, make sure you understand the various SD-Access fabric roles and components.

A *fabric role* is an SD-Access software construct that is running on physical hardware. A device can run a single role or multiple roles, with each having a specific responsibility. The SD-Access solution is made up of several components that work together. You need to fully understand the fabric roles for the design phase and choose the appropriate network devices for each role. The SD-Access fabric roles and components include the following:

- Control plane nodes
- Edge nodes
- Intermediate nodes
- Border nodes
- Fabric wireless LAN controllers (WLCs)
- Fabric-mode access points
- SD-Access embedded wireless
- Fabric in a Box
- Shared services

Figure 23.2 shows the SD-Access fabric roles.

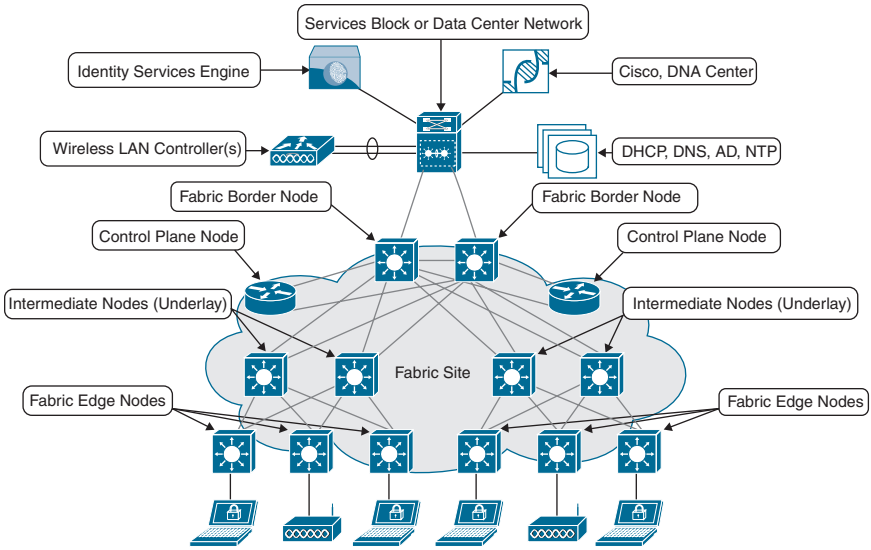



FIGURE 23.2 **SD-Access Fabric Roles**

The following sections examine these fabric roles and components.

# Control Plane Nodes

The SD-Access fabric control plane node is based on the use of LISP Map Server and Map Resolver on the same node. The database of the control plane node tracks all endpoints in the fabric site and associates the endpoints to fabric nodes, decoupling the endpoint IP address or MAC address from the location in the network.

The control plane node enables the following functions:

- **Host tracking database (HTDB):** The HTDB is a central repository of EID-to-RLOC bindings where the RLOC is the IP address of the Loopback 0 interface on a fabric node.
- **EID:** The EID is an address used for numbering or identifying an endpoint device in the network. In SD-Access, the supported addresses for EIDs are the MAC address, IPv4 address, and IPv6 address.
- **Map Server:** LISP Map Server receives endpoint registrations indicating the associated RLOC and uses this information to populate the HTDB.

- **Map Resolver:** LISP Map Resolver (MR) responds to queries from fabric devices requesting RLOC mapping information from the HTDB in the form of an EID-to-RLOC binding. This tells the requesting device to which fabric node an endpoint is connected and hence where to direct traffic.

A controller plane node must be either a Cisco switch or a router operating inside or outside the SD-Access fabric.

# Edge Nodes

The SD-Access fabric edge nodes are the equivalent of access layer switches in a traditional campus LAN design. The functionality of a fabric edge node is centered on the ingress and egress tunnel routers (xTR) in LISP. A Layer 3 routed access design is used for edge node implementation in an SD-Access deployment. An edge node enables the following functions:

- **Endpoint registration:** After the edge node detects an endpoint, that endpoint is added to a local database known as the EID table. Once the host is added to this local database, the edge node also issues a LISP map register message to notify the control plane node of the endpoint
- **Anycast Layer 3 gateway:** A common gateway (with IP and MAC addresses) is used at every edge node; it shares a common EID subnet and provides optimal forwarding and mobility across different RLOCs. This is represented as a switched virtual interface (SVI) with a hard-coded MAC address across all edge nodes on the edge node switches.
- **Mapping of user to the virtual network:** Endpoints connected to the edge nodes are placed into virtual networks by being assigned to VLANs associated with SVIs that are forwarding for VRF instances. Together, these make up the Layer 2 and Layer 3 LISP VNIs.
- **AAA authenticator:** The mapping of endpoints into VLANs can be done statically or dynamically by using an authentication server. The edge node is an essential part of the IEEE 802.1X port-based authentication process.
- **VXLAN encapsulation/decapsulation:** A packet or frame received from an endpoint is encapsulated in fabric VXLAN and forwarded across the overlay on the way to another edge node or border node. At the receiving end, it is decapsulated and sent to that endpoint.

The encapsulation and decapsulation of traffic enable the location of an endpoint to change without the endpoint having to change its address.

An edge node in SD-Access must be either a Cisco switch or router operating in the fabric overlay.

## Intermediate Nodes

Intermediate nodes are part of the Layer 3 network used for interconnections between border nodes and edge nodes. These interconnections are created in the global routing table on the devices and are known as the underlay network. Intermediate nodes do not have a requirement for VXLAN encapsulation/decapsulation, LISP control plane messaging support, or SGT awareness. An intermediate node must provide IP reachability and physical connectivity, and it must support the additional MTU requirement to support the larger-sized IP packets encapsulated with VXLAN information. The main idea here is that the intermediate nodes route and transport IP traffic between the devices operating in fabric roles.

## Border Nodes

A fabric border node serves as a gateway between the SD-Access fabric site and external networks to the fabric. A border node is responsible for network virtualization and SGT propagation from the fabric to the rest of the network. Border nodes implement the following functions:

- **Advertisement of EID subnets:** BGP is the routing protocol provisioned to advertise the coarse-aggregate endpoint prefix space outside the fabric.
- **Fabric site exit point:** An external border node is the gateway of last resort for the fabric edge nodes. It is implemented using LISP Proxy Tunnel Router (PxTR) functionality.
- **Network virtualization extension to the external world:** A border node can extend network virtualization from inside the fabric to outside the fabric by using VRF-Lite and VRF-aware routing protocols.
- **Policy mapping:** A border node maps SGT information from within the fabric to be appropriately maintained when exiting that fabric.
- **VXLAN encapsulation/decapsulation:** Packets and frames received from outside the fabric and destined for an endpoint inside the fabric are encapsulated in fabric VXLAN by the border node. Packets and frames sourced from inside the fabric and destined outside the fabric are decapsulated by the border node. This is similar to the behavior used by an edge node except that rather than being connected to endpoints, the border node connects a fabric site to a non-fabric network.

## Fabric Wireless LAN Controllers (WLCs)

In SD-Access Wireless, the Control and Provisioning of Wireless Access Points (CAPWAP) tunnels between the WLCs and APs are used for control traffic only. Data traffic from the wireless endpoints is tunneled to the first-hop fabric edge node, where security and policy can be applied at the edge, as they are for wired traffic. CAPWAP is discussed further in Chapter 20, "Wireless LAN Deployments." Cisco DNA Center configures wireless APs to reside within an overlay virtual network named INFRA_VN, which maps to the global routing table. This configuration eliminates the need for route leaking to establish connectivity between the WLCs and the APs.

## Fabric-Mode Access Points

The fabric-mode APs are the Cisco Wi-Fi 6 (802.11ax) and 802.11ac Wave 2 APs, and they are associated with the fabric WLCs that have been provisioned with one or more fabric-enabled SSIDs. They must be directly connected to a fabric edge node in the fabric site. The fabric APs establish a VXLAN tunnel to their fabric edge switch, where the wireless client traffic is terminated and placed on the wired network.

## SD-Access Embedded Wireless

Embedded Wireless using the Cisco Catalyst 9800 WLC is available for Catalyst 9000 Series switches as a software package on switches running in Install mode. When it is deployed in an SD-Access environment, CAPWAP tunnels are initiated on the APs and terminate on the Cisco Catalyst 9800 Embedded Wireless controller. The data plane uses VXLAN encapsulation for the overlay traffic between the APs and the fabric edge node.

The Catalyst 9800 Embedded Wireless controller for Catalyst 9000 Series switches is supported for SD-Access deployments with three topologies:

- Cisco Catalyst 9000 Series switches functioning as co-located border and control plane
- Cisco Catalyst 9000 Series switches functioning as edge nodes when the border and control plane nodes are on a routing platform
- Cisco Catalyst 9000 Series switches functioning as Fabric in a Box

# Fabric in a Box

Fabric in a Box is an SD-Access construct in which the border node, control plane node, and edge node are running on the same fabric node. This may be a single switch, a switch with hardware stacking, or a StackWise Virtual deployment.

# Shared Services

Some network services need to be shared across multiple virtual networks. Proper planning for deployment of the shared services is necessary to ensure that you correctly preserve network isolations when various virtual networks are accessing these services. A Fusion router assists in accomplishing this. Generally, all parts of the SD-Access fabric can be configured and managed using Cisco DNA Center. A Fusion device is outside the SD-Access fabric and is configured manually. A Fusion device enables VRF leaking across SD-Access fabric domains and enables host connectivity to shared services, such as DHCP, DNS, NTP, and so on.

These are some of the SD-Access shared services:

- **DHCP, DNS, IP Address Management (IPAM), and Active Directory (AD):** When infrastructure services support virtualized networks, they can be reused. Capabilities such as DHCP scoping selection criteria, multiple domains, and support for overlapping address space are some of the features required to extend the services beyond a single network.
- **Internet access:** The same set of Internet firewalls can be used for multiple virtual networks. In cases where unique firewall policies are needed for each virtual network, you can use a multi-context firewall.
- **IP voice/video collaboration services:** When unified communications devices are connected in multiple virtual networks, the call control signaling to the communications manager and the IP traffic between those devices needs to traverse multiple virtual networks in the infrastructure.
- **Servers and critical systems:** NTP servers, building management systems, network orchestrators, management appliances, databases, and other critical applications may be required for access by one or many virtual networks.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. What component in an SD-Access architecture is represented as a switched virtual interface (SVI) with a hard-coded MAC address across all edge nodes?
   - ❍ **A.** Map resolver
   - ❍ **B.** Anycast Layer 3 gateway
   - ❍ **C.** EID
   - ❍ **D.** PxTR

2. True or false: After an edge node detects an endpoint in the endpoint registration process, that endpoint is added to a local database known as the EID table.
   - ❍ **A.** True
   - ❍ **B.** False

## Answers

1. **B** is correct. An Anycast Layer 3 gateway is represented as an SVI with a hard-coded MAC address across all edge nodes.
2. **A** is correct. After an edge node detects an endpoint, that endpoint is added to a local database known as the EID table.

# Review Questions

1. True or false: Cisco SD-Access architecture offers open and standards-based APIs for management.
   - ❍ **A.** True
   - ❍ **B.** False

2. Which SD-Access operational plane deals with orchestration, assurance, and visibility in a Cisco SD-Access deployment?
   - ❍ **A.** Control plane
   - ❍ **B.** Data plane
   - ❍ **C.** Policy plane
   - ❍ **D.** Management plane

3. Fabric-mode APs connect to which of the following component in an SD-Access environment?
   - ❍ **A.** Fabric WLC
   - ❍ **B.** Edge node
   - ❍ **C.** Intermediate mode
   - ❍ **D.** Border node

# Answers to Review Questions

1. **A** is correct. Cisco SD-Access architecture offers open and standards-based APIs for management of a deployment.
2. **D** is correct. The management plane deals with orchestration, assurance, and visibility in a Cisco SD-Access deployment.
3. **B** is correct. Fabric mode APs must be directly connected to the fabric edge node switch in the fabric site.

# Further Reading

- ▶ *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*
- ▶ *Cisco Software-Defined Access, Cisco Secure Enterprise*
- ▶ *CCNP Enterprise Design ENSLD 300-420 Official Cert Guide: Designing Cisco Enterprise Networks*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers QoS.

*This page intentionally left blank*

CHAPTER 24

# QoS

**This chapter covers the following official ENCOR 350-401 exam objectives:**

- 1.6 Describe concepts of wired and wireless QoS
- 1.6.a QoS components
- 1.6.b QoS policy

This chapter looks at quality of service (QoS) as it relates to the measurement of transmission quality and service availability on a network. The first part of this chapter looks at the need for QoS. It surveys the causes of poor quality in network services and how to alleviate them with QoS tools. This chapter also covers several QoS models, components, and QoS policy. It examines the three models for implementing QoS in networks: best-effort, integrated services (IntServ), and differentiated services (DiffServ). This chapter also discusses classification to identify and assign IP traffic into different classes, as well as marking, which involves marking packets with priorities based on classification. In addition, this chapter looks at policing and shaping and how to enforce rate limiting by dropping, marking, or delaying excess IP traffic. This chapter also covers congestion management and avoidance to prioritize and protect IP traffic. Finally, this chapter looks at the application of QoS policy and QoS in WLAN environments.

**This chapter covers the following technology topics:**

- The Need for QoS
- QoS Models, Components, and Policy
- Congestion Management and Avoidance

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. What type of delay accounts for the time it takes for a router or switch to take a packet from an input interface and place it in the output queue of the output interface?
2. What QoS process is responsible for distinguishing one kind of traffic from another by examining the fields in the packet and determining which traffic is to be treated differently?

### Answers

1. Processing delay
2. Classification

# The Need for QoS

Modern networks play a critical part in the success of a lot of organizations. Networks transport a number of applications, including voice, video, and other delay-sensitive data. Thus, a network must provide predictable, measurable, and sometimes guaranteed services by handling bandwidth, delay, jitter, and loss parameters. Networks typically operate using best-effort delivery. This means that all traffic has equal priority and an equal chance of being delivered in a timely manner to its destination. When congestion occurs on the network, all traffic has an equal chance of being dropped.

The objective of QoS technologies is to make voice, video, and data convergence appear transparent to end users. QoS technologies allow different types of traffic to contend inequitably for network resources. Voice, video, and critical applications may be granted priority or preferential services on network devices so that the quality of critical applications does not degrade to the point of being useless.

When you implement QoS in a network environment, you can select specific network traffic and prioritize it according to its relative importance. You can then use traffic-management techniques to provide preferential treatment to specific types of traffic at the expense of other traffic types.

QoS implementation in a network makes network performance more predictable and bandwidth utilization more effective.

Without QoS, a network device offers best-effort service for each packet, regardless of the packet contents or size. The network device sends the packets without any assurance of reliability, delay, or throughput.

The following are specific features provided by QoS:

- Bandwidth guarantees
- Buffering capabilities and dropping mechanisms
- Traffic policing
- Traffic priorities
- Network traffic shaping

**ExamAlert**

For the ENCOR exam, make sure you know how packet loss, delay, jitter, and lack of bandwidth affects the transmission quality of the network.

To successfully implement QoS in a network, the network infrastructure should be designed to be highly available. Service availability is an essential foundation for QoS, and so is transmission quality. The following sections describe the factors that determine the transmission quality of the network:

- Packet loss
- Delay
- Jitter
- Lack of bandwidth

# Packet Loss

Packet loss is a relative measure of the number of packets that were not received compared to the total number of packets initially transmitted. Loss is generally a function of availability. Thus, a network must be designed with high availability in mind so that loss during periods of non-congestion is essentially zero. On the other hand, during periods of congestion, QoS mechanisms

can determine which packets would be appropriately dropped to alleviate the congestion. Packet loss can be addressed by implementing the following approaches:

- Increasing the speed of the link
- Implementing QoS congestion-avoidance and congestion-management techniques
- Implementing traffic policing to drop low-priority packets and allow high-priority traffic across
- Implementing traffic shaping to delay packets instead of dropping, as traffic may burst and exceed the capacity of an interface buffer at times

# Delay

Delay is the time it takes a packet to reach its destination after being transmitted from the sender. In the case of voice traffic, it is the amount of time it takes for sound to travel from the speaker's mouth to a listener's ear. Delay can be divided into fixed and variable types. Jitter, discussed next, is considered variable delay. The following types of delay are fixed:

- **Propagation delay:** Propagation delay is the time it takes for a packet to travel from a source endpoint to a destination endpoint over a fiber-optic or copper medium.
- **Serialization delay:** Serialization delay is the time it takes to put all the bits of a packet onto a link. It is a fixed value that depends on the speed of a link. That is, the higher the speed of the link, the lower the delay.
- **Processing delay:** *Processing delay* is the amount of time it takes for a router or switch to take a packet from an input interface and place the packet into the output queue of the output interface. A number of factors, including the following, can influence processing delay:
  - Packet switching mode (process switching, software CEF, or hardware CEF)
  - Router architecture (centralized or distributed)
  - CPU load
  - CPU speed on software-based platforms

## Jitter

Jitter (or delay variation) is a variable delay and is the difference in the end-to-end delay between packets. For example, if one packet requires 100 ms to traverse the network from the source to the destination and the following packet requires 125 ms to make the same trip, then the delay variation is 25 ms.

## Lack of Bandwidth

The bandwidth that is available on a packet data path from a source to a destination is equal to the lowest-bandwidth link's capacity in the data path. Once the lowest-bandwidth link's capacity is exceeded, congestion occurs, and packets are dropped. The simple fix would be to increase the link's bandwidth capacity, but technological or budgetary constraints may prevent this solution. Thus, you can use QoS mechanisms such as policing and queuing to prioritize traffic based on its importance. For example, time-sensitive and business-critical traffic like voice and video should be given priority forwarding, and the least important traffic should be given the remaining bandwidth.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Packet loss can be addressed by using which of the following approaches? (Choose all that apply.)
   - ❍ **A.** Increasing the speed of the link
   - ❍ **B.** Implementing QoS congestion-avoidance and congestion-management techniques
   - ❍ **C.** Implementing traffic policing
   - ❍ **D.** Avoiding use of traffic shaping

2. Which of the following is the accumulation of time it takes to put all the bits of a packet onto a link?
   - ❍ **A.** Propagation delay
   - ❍ **B.** Serialization delay
   - ❍ **C.** Processing delay
   - ❍ **D.** Jitter

# Answers

1. **A**, **B**, and **C** are correct. All of these help to avoid packet loss, including the use of traffic shaping, which is used to delay packets instead of drop them.
2. **B** is correct. Serialization delay is the time it takes to put all the bits of a packet onto a link. It is a fixed value that depends on the speed of a link.

# QoS Models and Components

A QoS service model describes a set of QoS capabilities. QoS service models differ in how they enable applications to send data and how the network attempts to deliver that data. You can use one service model for real-time applications, such as audio and video conferencing, and use another service model for file transfer and email applications. Cisco IOS supports three types of QoS service models: best effort, integrated, and differentiated services.

You should consider several factors before deciding which type of service to deploy in a network:

- Know the application or problem that you are trying to solve. Each model is appropriate for certain applications.
- Know the kind of capability that you want to allocate to your resources.
- Understand the cost–benefit analysis. The cost of implementing and deploying differentiated services, for example, is higher than the cost for best-effort service.

These are the different QoS models:

- **Best-effort service:** QoS is not actually implemented with the best-effort model. Rather, an application sends data whenever it needs, in any quantity, and without requesting permission or first informing the network. With best-effort service, packets are delivered without assurance of reliability, delay bounds, or throughput. Best-effort service is suitable for a wide range of applications, such as general file transfers and email.
- **Integrated service (IntServ):** In this model, the application requests a specific kind of service from the network (such as bandwidth reservation) before it sends data onto the network. The request is made by explicit signaling, where the application informs the network of its traffic profile and requests a specific kind of service that can meet its bandwidth and delay requirements. Once it gets a confirmation, the application is expected to send data onto the network within its described traffic profile. The network fulfills its commitment by maintaining a per-flow state and then performing packet classification, policing, and intelligent queuing based on that state.

  For IntServ, Cisco IOS QoS includes the following features that provide controlled load service, which is a kind of integrated service:

- **Resource Reservation Protocol (RSVP):** RSVP can be used by applications to signal their QoS requirements to the router.
- **Intelligent queuing mechanisms:** Mechanisms can be used with RSVP to provide the following kinds of services:
  - **Guaranteed rate service:** Applications can reserve bandwidth to meet their requirements.
  - **Controlled load service:** Applications can have low delay and high throughput even during times of network congestion.

- **Differentiated service (DiffServ):** With this model, it is possible to satisfy different QoS requirements. Unlike in the IntServ model, an application that is using DiffServ does not explicitly signal the router before sending data. For DiffServ, the network tries to deliver a particular kind of service based on the QoS specified by each packet. The specification occurs in a number of ways, such as by using the IP precedence bit settings in the IP packets or source and destination addresses. The network uses the QoS specification to classify, mark, shape, and police traffic and to perform intelligent queuing.

  The DiffServ model is the most commonly used QoS implementation model. It is used for a number of mission-critical applications and for providing end-to-end QoS. Cisco IOS QoS includes the following features that support the DiffServ model:

  - **Committed access rate (CAR):** The CAR makes it possible to perform metering and policing of traffic, providing bandwidth management.
  - **Intelligent queuing schemes:** Schemes such as weighted random early detection (WRED) and weighted fair queuing (WFQ) and their equivalent features can be used alongside CAR to deliver differentiated services.

Next, let's look at a number of QoS components that are necessary to implement any QoS mechanisms.

# Classification and Marking

Before applying any QoS mechanisms, you first need to identify and categorize IP traffic into different classes, based on your requirements. Network devices use classification to identify IP traffic as being part of a specific class. Once you have classified IP traffic, you need to use marking to mark individual packets so the network devices can apply QoS mechanisms to those packets as they move across the network.

Classification is the process of distinguishing one kind of traffic from another by examining the fields in the packet and determining which traffic is to be treated differently. During classification, the device performs a lookup and assigns a QoS label to the packet. The QoS label identifies all QoS actions that need to be performed on the packet and pinpoints the queue to which the packet is sent.

After classification, marking tools can set an attribute of a frame or packet to a specific value. Such marking (or remarking) establishes a trust boundary that scheduling tools later depend on in QoS implementation.

Classification and marking tools set the trust boundary by examining any of the following:

- **Layer 2 parameters:** 802.1Q class of service (CoS) bits and Multiprotocol Label Switching experimental values (MPLS EXP)
- **Layer 3 parameters:** IP precedence (IPP), differentiated services code point (DSCP), IP explicit congestion notification (ECN), and source/destination IP address
- **Layer 4 parameters:** Layer 4 protocol (TCP/UDP) and source/destination ports
- **Layer 7 parameters:** Application signatures via Network Based Application Recognition (NBAR)

NBAR is a Cisco-proprietary technology that identifies application layer protocols by matching them against a Packet Description Language Module (PDLM), which is basically an application signature. The NBAR deep-packet classification engine examines the data payload of stateless protocols against PDLMs. The NBAR operation is dependent on Cisco Express Forwarding (CEF), which performs deep-packet classification only on the first packet of a flow; the remainder of the packets belonging to the flow are then CEF switched.

**ExamAlert**

For the ENCOR exam, make sure you understand how marking is done at Layers 2 and 3.

Marking is done at either Layer 2 or Layer 3, using the following fields:

- **802.1Q/p class of service (CoS):** Ethernet frames can be marked at Layer 2 with their relative importance by setting the 802.1p user priority bits of the 802.1Q header. Only 3 bits are available for 802.1p marking. Therefore, only eight classes of service (0 through 7) can be marked on Layer 2 Ethernet frames.
- **IP type of service (ToS) byte:** Layer 2 media often changes as packets traverse from source to destination, so a more ubiquitous classification occurs at Layer 3. The second byte in an IPv4 packet is the ToS byte. The first 3 bits of the ToS byte are the IPP bits. These first 3 bits combined with the next 3 bits are known collectively as the DSCP bits.

Figure 24.1 shows the QoS classification layers in frames and packets.

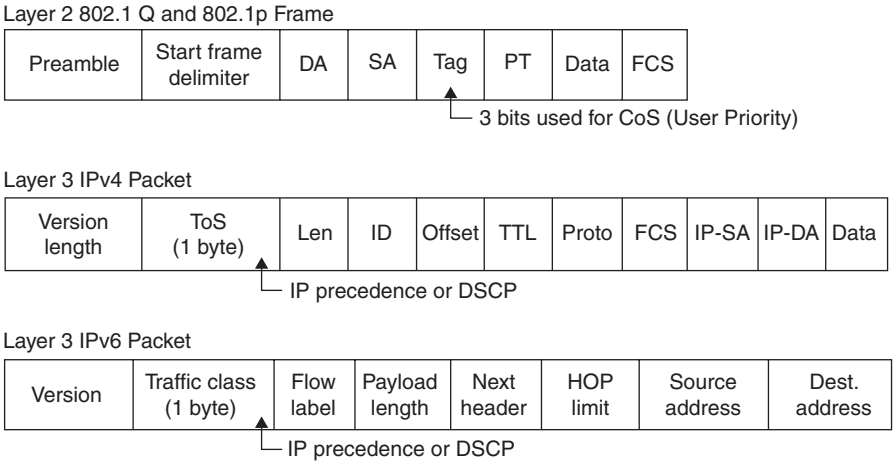



FIGURE 24.1 **QoS Classification Layers in Frames and Packets**

The IP precedence bits, like the 802.1p CoS bits, allow for only the following eight marking values (0 through 7):

- IPP values 6 and 7 are generally reserved for network control traffic, such as routing.
- IPP value 5 is recommended for voice.

- IPP value 4 is shared by videoconferencing and streaming video.
- IPP value 3 is for voice control.
- IPP values 1 and 2 can be used for data applications.
- IPP value 0 is the default marking value.

# DSCPs and Per-Hop Behaviors (PHBs)

DSCP values can be expressed in numeric form or by special standards-based names called PHBs. There are four broad classes of DSCP PHB markings:

- Best effort (BE or DSCP 0)
- RFC 2474 class selectors (CS1 through CS7, which are identical to and backward compatible with IPP values 1 through 7)
- RFC 2597 assured forwarding PHBs (AF*xy*)
- RFC 3268 Expedited Forwarding (EF)

There are four Assured Forwarding (AF) classes, each beginning with "AF" followed by two numbers (AF1, AF2, AF3, and AF4). The first number corresponds to the DiffServ class of the AF group and can range from 1 through 4. The second number refers to the level of drop precedence within each AF class and can range from 1 (lowest drop precedence) to 3 (highest drop precedence).

The AF name (AF*xy*) is composed of the AF IP precedence value in decimal (*x*) and the drop precedence value in decimal (*y*). For example, AF41 is a combination of IP precedence 4 and drop precedence 1.

IP Explicit congestion notification (IP ECN), as defined in RFC 3168, makes use of the last 2 bits of the IP ToS byte, which are not used by the 6-bit DSCP markings.

# Policing and Shaping

After packets are classified and have a DSCP-based, CoS-based, or QoS-group label assigned to them, the policing and marking process can start. If a session uses more than the allocated bandwidth, traffic is dropped (policing). Therefore, retransmission is necessary. Policers do not delay the traffic; they only check and take action when necessary.

Policing involves creating a policer that specifies the bandwidth limits for the IP traffic. Packets that exceed the limits are considered *out of profile* or *nonconforming*. Each policer determines, on a packet-by-packet basis, whether a

packet is in or out of profile and specifies the actions on the packet. The actions that are carried out by the marker include the following:

- Passing through the packet without modification
- Dropping the packet
- Modifying (marking down) the assigned DSCP or CoS value of the packet and allowing the packet to pass through

Shapers are considered traffic-smoothing tools that work in conjunction with buffering mechanisms. Shapers buffer or queue the traffic. They do not drop traffic, so retransmission is not necessary; however, shapers introduce latency. They smooth out traffic so it does not exceed the configured rate.

Shaping involves imposing a maximum rate of traffic while regulating the traffic rate so that downstream routers and switches are not subject to congestion. Shaping in the most basic form limits the traffic sent from a physical or logical interface. Shaping uses a buffer, which ensures that packets that do not have enough tokens are buffered rather than dropped immediately.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which of the following are considered QoS models? (Choose all that apply.)
   - ❍ **A.** Best-effort
   - ❍ **B.** NBAR
   - ❍ **C.** IntServ
   - ❍ **D.** DiffServ

2. True or false: Traffic shaping does not drop traffic, so retransmission is not necessary.
   - ❍ **A.** True
   - ❍ **B.** False

## Answers

1. **A**, **C**, and **D** are correct. Best-effort, IntServ, and DiffServ are all considered QoS implementation models.
2. **A** is correct. Traffic shaping does not drop traffic, so retransmission is not necessary. However, unlike policing, it introduces latency.

# Congestion Management and Congestion Avoidance

Now let's take a look at congestion management, or queuing, and congestion avoidance.

## Congestion Management (Queuing)

When you implement QoS, you can select specific network traffic, prioritize it according to its relative importance, and use congestion-management and congestion-avoidance techniques to provide preferential treatment. Network devices have buffers onboard that allow for scheduling higher-priority packets to exit sooner than lower-priority ones. This is commonly referred to as *queuing*. Queuing mechanisms are activated only when a device is experiencing congestion and are deactivated when the congestion no longer exists.

> **ExamAlert**
>
> For the ENCOR exam, be sure you know how to configure CBWFQ and LLQ.

You can configure the following types of queuing in Cisco IOS:

- **Class-based weighted fair queuing (CBWFQ):** CBWFQ provides bandwidth guarantees to given classes of traffic and fairness to discrete traffic flows within these traffic classes. With CBWFQ, you define traffic classes based on match criteria (including protocols, access control lists, and input interfaces). Packets fulfilling the match criteria for a class constitute the traffic for that class. A first-in, first-out (FIFO) queue is reserved for each class, and traffic belonging to a class is directed to the queue for that class. You then characterize a class by assigning it bandwidth, weight, and maximum packet limit. The bandwidth assigned to a class is the guaranteed bandwidth delivered to the class in the case of congestion.
- **Low-latency queuing (LLQ):** LLQ provides strict priority servicing and is intended for real-time applications, such as voice traffic. LLQ provides strict priority queuing for CBWFQ, reducing jitter in voice traffic. LLQ enables the use of a single, strict priority queue within CBWFQ at the class level, allowing you to direct traffic belonging to a class to the CBWFQ strict-priority queue.

# Congestion Avoidance

Congestion avoidance mechanisms are complementary to the queuing algorithms. They work by monitoring network traffic loads to anticipate congestion by dropping packets. Whereas queuing algorithms manage the front of a queue, congestion avoidance mechanisms manage the tail of the queue. The key to implementing the Cisco IOS congestion avoidance mechanism is WRED, which randomly drops packets as queues fill to capacity.

You can apply QoS features by using Modular QoS CLI (MQC). MQC allows you to create traffic policies and attaches those policies to interfaces. A traffic policy contains a traffic class and one or more QoS features. A traffic class is used to classify traffic, and the QoS features in a traffic policy determine how to treat the classified traffic. Defining the QoS policy involves these broad steps:

1. Define a traffic class by using the **class-map** command.
2. Create a traffic policy by using the **policy-map** command. A policy map contains a traffic class, and one or more QoS features are applied to that traffic class. The QoS features in the traffic policy determine how to treat the classified traffic.
3. Attach the traffic policy (via the policy map created in step 2) to the interface by using the **service-policy** command.

The other deployment option is AutoQoS, which Cisco created to simplify the deployment of QoS features. AutoQoS automatically matches traffic and assigns matched packets to QoS groups. This allows the output policy map to put specific QoS groups into specific queues, including into the priority queue. Using AutoQoS, you can quickly deploy and manage large-scale QoS deployments in your environment.

# Wireless QoS

Finally, to wrap up this chapter, let us now observe how a wireless LAN controller (WLC) uses QoS profiles for QoS implementation in wireless environments. Cisco Unified Wireless products support Wi-Fi Multimedia (WMM), a QoS system based on IEEE 802.11e that the Wi-Fi Alliance has published.

The IEEE 802.11e standard includes QoS features, user priorities, access categories, and user priority (UP)-to-DSCP mappings. The 802.11e standard introduced a 3-bit marking value in Layer 2 wireless frames known as UP. The value for the UP ranges from 0 to 7. Pairs of UP values are assigned to four access categories (AC), which equates to four distinct levels of service over a WLAN.

A WLAN can be configured with various default QoS profiles, where each of the QoS profiles is annotated with the typical use. Then clients can be assigned a QoS profile based on their identity through authentication, authorization, and accounting (AAA). On a Cisco WLC, the profiles can be configured as follows:

- **Bronze:** Background
- **Gold:** Video applications
- **Platinum:** Voice applications
- **Silver:** Best effort

Figure 24.2 shows the Cisco WLC QoS profiles.

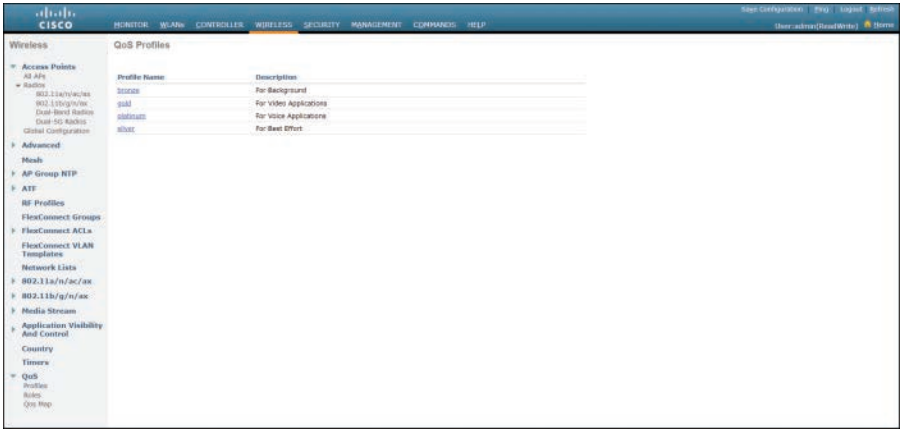



FIGURE 24.2 **Cisco WLC QoS Profiles**

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. True or false: Queuing mechanisms on a device are activated only when there is congestion and deactivated when congestion no longer exists.

   ❍ **A.** True

   ❍ **B.** False

2. True or false: QoS features in a traffic policy determine how to treat the classified traffic.
   - ❍ **A.** True
   - ❍ **B.** False

# Answers

1. **A** is correct. Queuing mechanisms are activated only when a device is experiencing congestion and deactivated when the congestion no longer exists.
2. **A** is correct. A traffic class is used to classify traffic, and the QoS features in a traffic policy determine how to treat the classified traffic.

# Review Questions

1. Which of the following is a variable delay and is the difference in the end-to-end delay between packets?
   - ❍ **A.** Propagation delay
   - ❍ **B.** Serialization delay
   - ❍ **C.** Processing delay
   - ❍ **D.** Jitter

2. Which of the following QoS implementation models is most commonly used?
   - ❍ **A.** Best-effort
   - ❍ **B.** Assured forwarding
   - ❍ **C.** IntServ
   - ❍ **D.** DiffServ

3. Which of the following queuing methods provides strict priority servicing and is intended for real-time applications such as voice traffic?
   - ❍ **A.** CBWFQ
   - ❍ **B.** LLQ
   - ❍ **C.** WRED
   - ❍ **D.** DWRED

# Answers to Review Questions

1. **D** is correct. Jitter (or delay variation) is a variable delay and is the difference in the end-to-end delay between packets.
2. **D** is correct. The differentiated service model is the most commonly used QoS implementation model.
3. **B** is correct. Low-latency queuing (LLQ) provides strict priority servicing and is intended for real-time applications such as voice traffic.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*
- *End-to-End QoS Network Design: Quality of Service for Rich-Media & Cloud Networks, Second Edition*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers hardware and software switching mechanisms.

CHAPTER 25

# Switching

**This chapter covers the following official ENCOR 350-401 exam objectives:**

- 1.7 Differentiate hardware and software switching mechanisms
- 1.7.a Process and CEF
- 1.7.b MAC address table and TCAM
- 1.7.c FIB vs. RIB

This chapter reinforces the concepts related to how switches and routers forward data at their respective layers. It examines how switches forward traffic from a Layer 2 perspective and how routers/multilayer switches (MLSs) forward traffic from a Layer 3 perspective. This chapter starts by reviewing basic network fundamentals and then covers the forwarding architectures that switches use, including process switching, fast switching, and Cisco Express Forwarding (CEF). This chapter also reviews the different memory tables used in switching, including content-addressable memory (CAM) and ternary content-addressable memory (TCAM). Finally, this chapter examines the use of Switch Database Manager (SDM) templates for managing switches' internal resources.

**This chapter covers the following technology topics:**

- Traffic Forwarding Basics
- Forwarding Architectures

## Cram Saver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. In Layer 2 forwarding, what does a switch do after CAM and TCAM table lookups?
2. Which switching method uses the general-purpose CPU to handle all packet switching?
3. What are the two components of Cisco Express Forwarding (CEF)?
4. What command is used for changing the SDM template on a Cisco switch?

### Answers

1. Frames are placed in the appropriate egress queue on the outbound switch ports of the switch.
2. Process switching
3. Forwarding information base (FIB) and adjacency table
4. **sdm prefer**

# Traffic Forwarding Basics

From a Layer 2 perspective, switches try to address the limitation of operating an Ethernet network using hubs, where hosts are connected to a single broadcast and collision domain. An Ethernet switch provides isolation from other connected hosts, offering several improvements over the use of hubs:

- The collision domain scope is limited; on each switch port, the scope of the collision domain is limited to the port itself and the device connected to that port.
- Each switch port offers dedicated bandwidth, so bandwidth does not need to be shared.
- The host can operate in full-duplex mode because there is no longer any contention on the media.

- Errors in frames are no longer propagated because each frame received on a switch port is checked for errors.
- It is possible to limit broadcast traffic by segmenting LANs into VLANs using grouping of switch ports.

When frames arrive on a switch port, they are placed in an ingress queue. These queues with frames to be forwarded can have different priorities, or service levels. The switch ports can be configured to have more important frames processed before less important frames so that time-critical traffic does not get lost in the midst of incoming traffic.

As ingress queues are serviced and a switch figures out where to forward traffic, it also needs to figure out which egress switch port it needs to forward traffic out of and what forwarding policies to use. These decisions are made by independent portions of the switching hardware:

- **Layer 2 forwarding table:** The frame destination MAC address is used as an index to the content-addressable memory (CAM) table. If the address is found, the egress switch port and VLAN are read from the table. If the address is not found, the frame is marked for flooding out of every switch port in that particular VLAN, except the receiving port.
- **Security access control lists (ACLs):** ACLs can be used to identify frames according to their MAC addresses, IP addresses, protocols, and Layer 4 port numbers. Ternary content-addressable memory (TCAM) contains ACLs in a compiled format to allow decisions to be made on whether to forward a frame in a single lookup. (TCAM operations are covered later in this chapter.)
- **Quality of service (QoS) ACLs:** ACLs can classify incoming frames according to QoS parameters to control or police the traffic flow rate and mark QoS parameters in outbound frames. TCAM is also used to make these decisions in a single table lookup.

After CAM and TCAM table lookups, frames are placed in the appropriate egress queue on the outbound switch ports of a switch. Like the ingress queue, the egress queue is serviced according to importance or time criticality. (The next section looks at traffic forwarding from a Layer 3 or multilayer switching perspective.)

Figure 25.1 shows a Layer 2 switch and the decision process it uses to forward each frame.

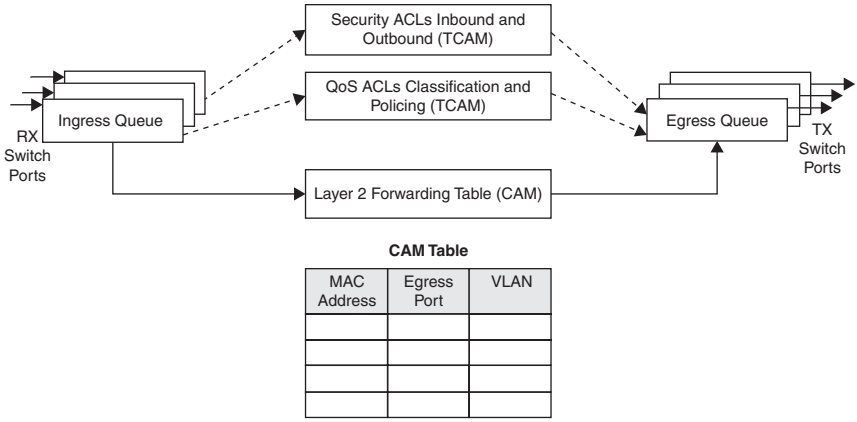



FIGURE 25.1 **Operations in a Layer 2 Switch**

The path a Layer 3 packet follows through an MLS is similar to that of a Layer 2 switch. However, some means of making a Layer 3 forwarding decision must be added.

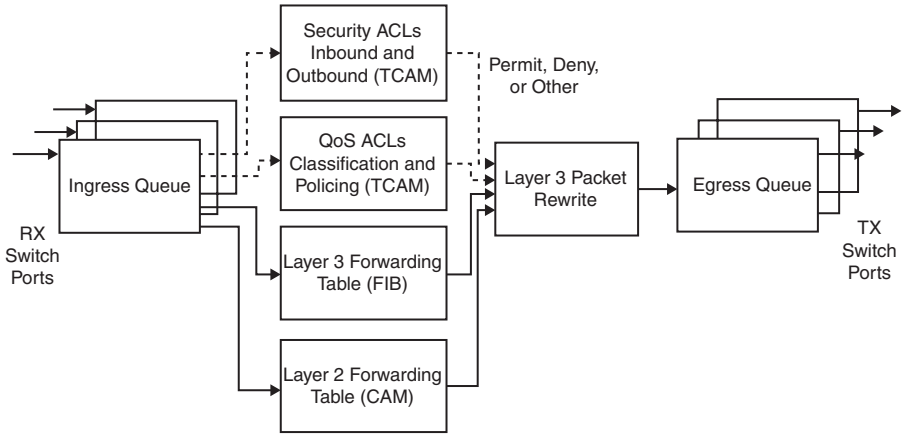
Each packet is inspected for both Layer 2 and Layer 3 addresses after being pulled off an ingress queue. However, the decision about where to forward the packet is based on two tables. As with Layer 2 switching, all the multilayer switching decisions are made simultaneously in hardware:

- **Layer 2 forwarding table:** The destination MAC address is used as an index to the CAM table. If there are Layer 3 packets to be forwarded, the destination MAC address is the address of the Layer 3 port on the switch. In this case, the CAM table results are only used to decide that the frame should be processed at Layer 3.
- **Layer 3 forwarding table:** The forwarding information base (FIB) is consulted using the destination IP address as an index. The longest match is found in the table using both the address and the mask, and the resulting next-hop Layer 3 address is obtained. The FIB already has the next-hop entry Layer 2 MAC address and the egress port and VLAN, so no further processing is needed at this point.
- **Security ACLs:** Inbound and outbound access lists are compiled into TCAM entries to make decisions about whether to forward a packet; this can be done as a single table lookup.
- **QoS ACLs:** Packet classification, policing, and marking can all be performed as a single table lookup in the QoS TCAM.

As with Layer 2 switching, with Layer 3 switching, the packet is placed in an egress queue on the appropriate egress switch port. As a router would, the MLS at this point already has the next-hop destination from the FIB table (from the preceding list). The Layer 3 address identifies the next hop and finds its Layer 2 address; at this point, only the Layer 2 address would be used to forward Layer 2 frames.

The next-hop Layer 2 address is used in the frame in place of the original destination Layer 2 address. The Layer 2 source address in the frame now becomes the address of the MLS before being sent on to the next hop. At this stage, the Time-to-Live (TTL) value is decremented by one. Because the content of the Layer 3 packet (that is, the TTL value) has changed, the Layer 3 header checksum must be recalculated. The Layer 2 checksum also needs to be recalculated because at this point both the Layer 2 and the Layer 3 contents have changed. In other words, the entire Ethernet frame must be rewritten before it goes into the egress queue, and this rewriting can be done efficiently in hardware.

Figure 25.2 shows a multilayer switch and the decision processes it uses to forward each frame.




**CAM Table**

| MAC Address | Egress Port | VLAN |
|---|---|---|
| | | |
| | | |
| | | |

**FIB Table**

| IP Address | Next-Hop IP Address | Next-Hop MAC Address | Egress Port |
|---|---|---|---|
| | | | |
| | | | |
| | | | |

FIGURE 25.2 **Operations in a Multilayer Switch**

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. What table do Cisco switches use to determine which port a frame should be out?
   - ❍ **A.** ARP table
   - ❍ **B.** Routing table
   - ❍ **C.** MAC address table
   - ❍ **D.** VLAN table

2. What information does a Layer 2 device use to forward network traffic?
   - ❍ **A.** Source IP address
   - ❍ **B.** Destination IP address
   - ❍ **C.** Source MAC address
   - ❍ **D.** Destination MAC address

3. Which buffer is a frame placed in after a forwarding decision has been made?
   - ❍ **A.** Ingress queue
   - ❍ **B.** Egress queue
   - ❍ **C.** RIB
   - ❍ **D.** FIB

## Answers

1. **C** is correct. The MAC address table records source MAC addresses as well as the port and VLAN on which each address was learned. As a result, this table can be used to determine the port a frame should be forwarded out.
2. **D** is correct. A switch uses the destination MAC address in a frame to determine which port the frame needs to be forwarded out of.
3. **B** is correct. After CAM and TCAM table lookups, frames are placed in the appropriate egress queue on the outbound switch ports.

# Forwarding Architectures

Over the years, advances in technologies have facilitated the streamlining of the process used to forward IP packets. Previously, packet forwarding started with process switching. However, routers do not need to remove and add the Layer 2 addresses; they just need to rewrite the addresses. IP packet forwarding is a faster process for receiving an IP packet on an input interface and deciding to forward it out of an output interface or drop it. As network devices evolved, Cisco began using Cisco Express Forwarding (CEF) to optimize a router's switching process to facilitate forwarding of large numbers of packets.

The three switching mechanisms that Cisco devices use to make forwarding decisions are process switching, fast switching, and CEF.

**ExamAlert**

Before taking the ENCOR exam, make sure you understand the three switching mechanisms used by Cisco devices when making forwarding decisions.

## Process Switching

Process switching, or software switching, is a switching mechanism in which the device's general-purpose CPU handles packet switching. Process switching is the fallback for Cisco Expressing Forwarding (CEF) switching when IP packets need to be punted to the general-purpose CPU when they cannot be CEF switched.

Process switching is the slowest method of packet switching because it requires the general-purpose CPU of the devices to be directly involved in processing packets that come into and go out of a device. This processing by the general-purpose CPU of the device adds delay to the packet. On modern Cisco devices, process switching is only used when handling certain types of traffic.

The types of traffic that require software handling and thus processing switching include the following:

- Packets that have IP options, since these packets are too complex for the hardware to handle
- Packets that are sourced or destined to the router itself
- Packets that require extra information that is not known (for example, from ARP tables)

# Fast Switching

After process switching, fast switching was Cisco's next evolution in packet switching technology. It works by implementing a high-speed cache to accelerate the speed of packet processing. In fast switching, the first packet is copied to the system buffer. The device then looks up the Layer 3 address in the routing table and initializes the fast-switching cache. The frame is then rewritten with the destination address, and the frame is sent to the outgoing interface that services that destination. The route processor also runs a cyclical redundancy check (CRC). When the subsequent frame arrives, the destination is found in the fast-switching cache. Furthermore, all subsequent packets to that same destination are sent by the same switching path.

# Cisco Express Forwarding (CEF)

CEF was Cisco's next evolution of packet switching technology, and it is used on most Cisco devices by default. CEF switching is implemented as both software-based CEF and hardware-based CEF. In software-based routers, the general-purpose CPU manages all operations, including software CEF switching. In more advanced hardware-based platforms, routers do CEF switching using forwarding engines that are implemented in specialized application-specific integrated circuits (ASICs), ternary content addressable memory (TCAM), and network processing units (NPUs) for hardware CEF switching.

The following are some of the main benefits of CEF:

- **Improved performance:** CEF is less CPU intensive than fast switching route caching. As a result, more CPU resources can be dedicated to Layer 3 services, such as QoS.
- **Scalability:** CEF offers full switching capacity at each line card when distributed CEF is active.
- **Resilience:** In dynamic networks, fast-switched cache entries are frequently invalidated by routing changes, which can cause fallback to process switching rather than to fast switching using route caching. Since forwarding information base (FIB) tables contain all known routes that exist in the routing table, there is no need for route cache maintenance, fast switching, and process switching forwarding.

**ExamAlert**

Before taking the ENCOR exam, make sure you understand the components of CEF, including the makeup of the FIB and adjacency tables.

## Components of CEF

The data stored in the route cache is stored in several data structures for CEF switching. The data structures provide optimum lookup for efficient packet forwarding. The components of CEF forwarding operations are as follows:

- **Forwarding information base (FIB):** The FIB, which is used to make IP destination prefix-based switching decisions, is built from the RIB. It is similar to a routing table, and routers use it to make destination-based switching decisions during CEF operation. As changes on the network occur, the FIB is updated to reflect all routes known at the time. Because there is a one-to-one correlation between the FIB entries and the routing table entries, the FIB contains all the known routes. This eliminates the need for the route cache maintenance that is associated with switching methods such as fast switching.
- **Adjacency tables:** A node is said to be adjacent to another node if it can be reached with a single hop over the link layer (Layer 2). CEF stores the forwarding information (outbound interface and MAC header rewrite) for the adjacent node in a data structure known as the *adjacency table*. CEF uses adjacency tables to prepend Layer 2 addressing information to packets. The adjacency tables maintain Layer 2 next-hop addresses for all FIB entries and are built from the information found in the ARP cache.

The commands **show ip cef** and **show adjacency** can be used to verify the FIB and adjacency tables, respectively.

The separation of the reachability information (in the CEF table) and the forwarding information (in the adjacency table) provides the following benefits:

- The adjacency table can be built separately from the CEF table, basically allowing both to be built without the packet being process switched.
- The MAC header rewrite used to forward a packet is not stored in cache entries; therefore, changes in a MAC header rewrite string do not require validation of cache entries.

Example 25.1 shows sample output from the **show ip cef** and **show adjacency** commands on a Catalyst 9300 switch using CEF forwarding. The **show ip cef** command output shows the prefix, next hop, and interface, and the **show adjacency** output shows the interface and address.

EXAMPLE 25.1 **Output of the show ip cef and show adjacency Commands**

```
SW1# show ip cef
Prefix               Next Hop             Interface
0.0.0.0/0            192.168.10.12        Vlan10
0.0.0.0/8            drop
0.0.0.0/32           receive
10.10.1.0/24         192.168.10.12        Vlan10
192.168.50.0/24      attached             Vlan50
192.168.50.0/32      receive              Vlan50
192.168.110.0/24     192.168.50.2         Vlan50
192.168.120.0/24     attached             Vlan120
<...output omitted...>

SW1# show adjacency
Protocol Interface                 Address
IP       Vlan1                     192.168.1.7(8)
IP       Vlan50                    192.168.50.2(10)
IP       Vlan50                    192.168.50.10(8)
IP       Vlan120                   192.168.120.10(8)
<...output omitted...>
```

# CEF Modes of Operation

CEF switching can occur at two different locations on a switch. Based on the location where CEF operates, it can be used in either of these modes:

- **Centralized CEF mode:** This mode is used when line cards are not available for CEF forwarding, when there is a need to run features that are not compatible with distributed CEF, or when you are running a non-distributed platform. When centralized CEF mode is enabled, the CEF FIB and adjacency tables reside on the route processor (RP), and the RP performs the express forwarding.
- **Distributed CEF (dCEF) mode:** This mode is used on certain platforms where processing tasks are spread across two or more line cards to achieve scalability. After the distributed CEF mode of operation is enabled, line cards maintain identical copies of the FIB and adjacency tables. Distributed CEF increases system performance because line cards perform the express forwarding between port adapters, eliminating RP involvement in the switching operation. Distributed CEF also uses an interprocess communication (IPC) mechanism to synchronize FIB tables and adjacency tables on the RP and line cards.

The command **ip cef** is used to enable CEF switching, and **no ip cef** is used to disable CEF switching. You can determine whether a switch is operating in centralized or distributed CEF mode by using the **show ip cef summary** command.

Example 25.2 shows how to verify different CEF modes.

EXAMPLE 25.2 **Verifying CEF Modes**

```
SW1#
SW1# show ip cef summary
IPv4 CEF is enabled for distributed and running
VRF Default
 606 prefixes (605/1 fwd/non-fwd)
 Table id 0x0
 Database epoch:        3 (606 entries at this epoch)
SW1#

SW2#
SW2# show ip cef summary
IPv4 CEF is enabled and running
VRF Default
 8 prefixes (7/1 fwd/non-fwd)
 Table id 0x0
 Database epoch:        0 (8 entries at this epoch)
SW2#
```

# Tables Used in Switching

Cisco Catalyst switches use two types of tables in the switching process: content-addressable memory (CAM) table and the ternary content-addressable memory (TCAM) table. These tables are kept in high-speed memory so that many fields within a frame or packet can be compared in parallel.

## Content-Addressable Memory (CAM) Table

As frames arrive on a switch port, the CAM table records the source MAC address learned on the port. The MAC address is recorded along with the port, VLAN, and a timestamp. If the MAC address learned on a port moves to a different port, the MAC address and timestamp are recorded for the most recent port. At that point, the previous entry is deleted. If the MAC address is already present in the CAM table for the correct arrival port, only the timestamp is updated.

Switches generally have large CAM tables to look up addresses. However, if there is not enough space to hold every possible entry, stale entries (that is, addresses that have not been heard from for a period of time) are aged out. By default, idle CAM table entries are kept for 300 seconds before they are removed. This default setting can be changed to suit the environment.

The CAM table manages duplicate entries using a system of purging. For example, if a device's MAC address that was learned by a switch moves to a different port, the CAM table ages out the entry after 300 seconds. To avoid any duplicate entries, a switch purges an existing entry for a MAC address that has just been learned on a different switch port. Because MAC addresses are unique and should not exist on multiple ports, purging is safe. However, if a MAC address is being learned on multiple ports, this would generate an error because it indicates that the MAC address is "flapping" between interfaces.

You can display the CAM table by using the **show mac address-table** command. To verify the current aging time, you use the command **show mac address-table aging-time**. You change the default aging time by using the command **mac address-table aging-time** *seconds*.

Example 25.3 shows how to verify the contents of the CAM table, its aging time, and the aging time configuration.

EXAMPLE 25.3 **Verifying the CAM Table**

```
SW1#
SW1# show mac address-table
          Mac Address Table
-------------------------------------------

Vlan    Mac Address       Type        Ports
----    -----------       --------    -----
   1     5254.000f.d346    DYNAMIC     Gi0/2
   10    5254.0018.6cbd    DYNAMIC     Gi0/3
   10    5254.001a.37c2    DYNAMIC     Gi0/4
   20    5254.001c.827f    DYNAMIC     Gi0/5
   20    5254.001d.657b    DYNAMIC     Gi0/6
Total Mac Addresses for this criterion: 5
SW1#

SW1#
SW1# show mac address-table aging-time
Global Aging Time:  300
Vlan    Aging Time
----    ----------
SW1#
```

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# mac address-table aging-time ?
  <0-0>         Enter 0 to disable aging
  <10-1000000>  Aging time in seconds

SW1(config)# exit
SW1#
```

## Ternary Content-Addressable Memory (TCAM)

As mentioned earlier, switches use specialized hardware to house the MAC address table, QoS lookup data, and ACLs. To house the MAC address table, switches use CAM, and for housing ACL and QoS tables, they use TCAM. Both CAM and TCAM provide high-speed access and allow for line-rate switching performance. However, CAM only matches on ones and zeros (binary), and this limitation created the need for TCAM.

TCAM can match on a third state, which is any value. TCAM can provide three results: 0, 1, or X (do not care), which is a ternary combination. This offers more flexibility in searching than does the binary CAM. TCAM is useful in building tables for searching on longest matches, such as routing tables organized by IP prefixes. TCAM is not only an extension of the CAM architecture but enhanced to store information generally associated with upper-layer protocols, such as ACLs and QoS. For this reason, most switches contain multiple TCAM tables so that both inbound and outbound ACL and QoS ACLs can be evaluated at the same time or in parallel with Layer 2 or Layer 3 forwarding decisions.

There are two components of TCAM operation:

- **Feature Manager (FM):** After an ACL is created, the FM software compiles or merges the access control entries (ACEs) into entries in the TCAM table; the TCAM can then be consulted at full frame forwarding speed.
- **Switching Database Manager (SDM):** SDM partitions the TCAM on a switch into different areas for different functions.

SDM templates are used to configure system resources in a switch to optimize support for specific features, depending on how the switch will be used in the network environment. The switch SDM templates allocate system hardware resources for different use cases. For example, the default template can be used

to balance resources, and the access template can be used to manage a large number of ACLs.

To allocate TCAM resources for different usage types, the SDM templates prioritize system resources to optimize support for certain features. SDM templates can be selected and optimized for the following features:

- **Access:** The access template maximizes systems resources for ACLs to accommodate a large number of ACLs.
- **Default:** The default template gives balance to all functions.
- **Routing:** The routing template maximizes resources for unicast routing and is typically required for a router or an aggregator in the center of a network.
- **VLANs:** The VLAN template disables routing and supports the maximum number of MAC addresses; this template would typically be selected for Layer 2 switches.
- **Dual IPv4 and IPv6:** The dual IPv4 and IPv6 template allows a switch to be used in dual-stack environments.
- **Advanced:** The advanced template maximizes system resources for features like NetFlow, multicast groups, security ACEs, and QoS ACEs.

You can change the SDM template by using the **sdm prefer** command.

Example 25.4 shows the options for setting the SDM template.

EXAMPLE 25.4 **Changing the SDM Template**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# sdm prefer ?
  advanced  Advanced Template
  vlan      VLAN Template
SW1(config)# exit
SW1#

SW2#
SW2# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW2(config)# sdm prefer ?
  access              Access bias
  default             Default bias
  dual-ipv4-and-ipv6  Support both IPv4 and IPv6
```

```
  routing             Unicast bias
  vlan                VLAN bias
SW2(config)# exit
SW2#
```

Example 25.5 illustrates the use of the **show sdm prefer** command to view the current SDM templates on two switches: the advanced template and the default template.

EXAMPLE 25.5 **Verifying the SDM Template**

```
SW1# show sdm prefer
Showing SDM Template Info

This is the Advanced template.
  Number of VLANs:                                4094
  Unicast MAC addresses:                          32768
  Overflow Unicast MAC addresses:                 512
  L2 Multicast entries:                           4096
  Overflow L2 Multicast entries:                  512
  L3 Multicast entries:                           4096
  Overflow L3 Multicast entries:                  512
  Directly connected routes:                      16384
  Indirect routes:                                7168
  STP Instances:                                  4096
  Security Access Control Entries:                3072
  QoS Access Control Entries:                     2560
  Policy Based Routing ACEs:                      1024
  Netflow ACEs:                                   768
  Flow SPAN ACEs:                                 512
  Tunnels:                                        256
  LISP Instance Mapping Entries:                  256
  Control Plane Entries:                          512
  Input Netflow flows:                            8192
  Output Netflow flows:                           16384
  SGT/DGT (or) MPLS VPN entries:                  4096
  SGT/DGT (or) MPLS VPN Overflow entries:         512
  Wired clients:                                  2048
  MACSec SPD Entries:                             256
  MPLS L3 VPN VRF:                                127
  MPLS Labels:                                    2048
  MPLS L3 VPN Routes VRF Mode:                    7168
  MPLS L3 VPN Routes Prefix Mode:                 3072
  MVPN MDT Tunnels:                               256
  L2 VPN EOMPLS Attachment Circuit:               256
  MAX VPLS Bridge Domains :                       64
  MAX VPLS Peers Per Bridge Domain:               8
  MAX VPLS/VPWS Pseudowires :                     256
```

```
These numbers are typical for L2 and IPv4 features.
Some features such as IPv6, use up double the entry size;
so only half as many entries can be created.
* values can be modified by sdm cli.
SW1#
SW2#
SW2# show sdm prefer
 The current template is "desktop default" template.
 The selected template optimizes the resources in
 the switch to support this level of features for
 8 routed interfaces and 1024 VLANs.

  number of unicast mac addresses:                  6K
  number of IPv4 IGMP groups + multicast routes:    1K
  number of IPv4 unicast routes:                    8K
    number of directly-connected IPv4 hosts:        6K
    number of indirect IPv4 routes:                 2K
  number of IPv6 multicast groups:                  0
  number of directly-connected IPv6 addresses:      0
  number of indirect IPv6 unicast routes:           0
  number of IPv4 policy based routing aces:         0
  number of IPv4/MAC qos aces:                      0.5K
  number of IPv4/MAC security aces:                 1K
  number of IPv6 policy based routing aces:         0
  number of IPv6 qos aces:                          20
  number of IPv6 security aces:                     25
 SW2#
```

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. In most modern Cisco switches, which switching method is used by default?
   - ❍ **A.** CEF switching
   - ❍ **B.** Process switching
   - ❍ **C.** Fast switching
   - ❍ **D.** Cut-through switching

2. Which switching method uses the general-purpose CPU to process all packets?
   - ❍ **A.** CEF switching
   - ❍ **B.** Process switching
   - ❍ **C.** Fast switching
   - ❍ **D.** Cut-through switching

3. Which components form the data structure of CEF? (Choose two.)
   - ❍ **A.** CAM table
   - ❍ **B.** Forwarding information base
   - ❍ **C.** Routing information base
   - ❍ **D.** Adjacency table

# Answers

1. **A** is correct. CEF is the packet switching technology that is used on most Cisco devices by default.
2. **B** is correct. Process switching, or software switching, is a switching mechanism in which the device's general-purpose CPU handles packet switching.
3. **B** and **D** are correct. The two components that form the data structure of CEF are the forwarding information base (FIB) and the adjacency table.

# Review Questions

1. Which CEF mode of operation makes it possible to achieve higher scalability on the switching platform?
   - ❍ **A.** Centralized CEF
   - ❍ **B.** Distributed CEF
   - ❍ **C.** Process switching
   - ❍ **D.** Fast switching

2. Which switching method uses information gathered from tables, such as routing tables and Address Resolution Protocol (ARP) tables, to build hardware-based tables, such as forwarding information base (FIB) and adjacency tables?
   - ❍ **A.** CEF switching
   - ❍ **B.** Process switching
   - ❍ **C.** Fast switching
   - ❍ **D.** Cut-through switching

3. What does a switch do if a MAC address cannot be found in the CAM table?
   - ❍ **A.** Drops the frame
   - ❍ **B.** Generates an ARP request for that address
   - ❍ **C.** Floods the frame out all ports (except the receiving port)
   - ❍ **D.** Floods the frame out all ports (including the receiving port)

# Answers to Review Questions

1. **B** is correct. Distributed CEF is used on certain platforms where processing tasks are spread across two or more line cards to achieve scalability.
2. **A** is correct. CEF switching uses data structures known as FIB and adjacency tables that are built from information gathered from tables such as routing tables and ARP tables.
3. **C** is correct. If the address is not found, the frame is marked for flooding out of every switch port in that particular VLAN, except the receiving port.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers basic virtualization.

*This page intentionally left blank*

CHAPTER 26

# Basic Virtualization

**This chapter covers the following official ENCOR 350-401 exam objectives:**

- 2.1 Describe device virtualization technologies
- 2.1.a Hypervisor type 1 and 2
- 2.1.b Virtual machine
- 2.1.c Virtual switching

One technology that has been a significant game-changer in the IT industry in recent years is virtualization. Today, virtualization is found throughout enterprise campus and data center deployments. This chapter covers server virtualization, which involves running multiple virtual instances of an operating system and applications in a layer abstracted from the underlying hardware. It dives into hypervisors, virtual machines (VMs), virtual switching, and network virtualization. This chapter is divided into three sections. The first section provides an overview of virtualization. The next section looks at VMs and virtual switching. The final section covers network virtualization.

**This chapter covers the following technology topics:**

- Virtualization Overview
- Virtual Machines (VMs)
- Virtual Switching

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. What does consolidation ratio mean in the context of virtualization?
2. Which type of hypervisor is most commonly used in enterprise production deployments?
3. From the hypervisor perspective, how is a virtual disk represented?
4. In network virtualization, what is the terminology given to the physical enterprise network that supports multiple logical overlay networks?

### Answers

1. Consolidation ratio is the ratio of virtual servers to physical servers.
2. Type 1, or bare-metal hypervisor
3. A virtual disk is represented as a file on the host-attached storage of a hypervisor.
4. Underlay network

# Virtualization Overview

When engineers hear the term *virtualization*, they typically associate it with server virtualization. Cisco has been doing virtualization for years. One common virtualization technology Cisco has been using is virtual LANs (VLANs). In its most basic form, virtualization with VLANs allows for the creation of multiple virtual broadcast domains by moving specific switch ports into a particular VLAN and using a Layer 3 device to route between them.

One of the critical drivers for server virtualization is underuse of server hardware resources. The traditional method was to pair an operating system and a single application to a physical server for availability reasons; the goal was to prevent failure of one application running on a server from affecting another application or a compatibility issue causing unforeseen problems among servers. Although this approach provides good stability, this "one workload to one box" model is inefficient. An application may use only 5% to 20% of the server resources, resulting in underutilization and overprovisioning of resources. In addition, this approach does not scale well because increasing the number of

applications increases the number of physical servers needed, which increases the total cost of ownership (TCO).

Virtualization increases the overall efficiency of a physical server by maximizing the use of its resources. Virtualization allows for server consolidation because it allows you to run multiple virtual machines on one or more physical servers. It is not uncommon to run 20 to 30 virtual servers on a single physical server. The ratio of virtual servers to physical servers is known as the consolidation ratio; a higher consolidation ratio provides a higher return on investment (ROI) on the virtualization deployment. Running multiple workloads on highly configured x86 hardware can increase the utilization of the physical server hardware to as much as 80%.

Virtualization also increases the cost-effectiveness of operating the infrastructure because there are fewer physical servers to manage, less power is required, and less cabling and rack space are needed. In contrast, in a non-virtualized environment, the Windows or Linux operating system directly accesses all physical server resources. If a server has 8 CPU cores and 16 GB of memory, all will be available to the operating system and applications. Virtualization also speeds up deployment time for VMs/virtual appliances because it eliminates the time required to procure a physical server and handle racking, cabling, configuration, and so on. When you go down the path of server virtualization, you have a cluster or pool of resources ready to deploy to VMs/virtual appliances.

Besides using virtualization for production in private and public cloud deployments, you can use it in a test/development environment. Operating systems and applications can be cloned from a production virtual environment into a test environment to test operating system upgrades, patching, or application updates without causing potentially disastrous effects in the production environment. After an upgrade or update has been tested to satisfaction in a non-production environment, it can be deployed in the production environment. An engineer can then do the production rollout with confidence and less uncertainty of the outcome, knowing that the solution was tested in an identical non-production environment.

In addition to virtualizing servers, you can also virtualize networks. Network virtualization makes it possible for a single physical network topology to be used by different virtual networks that have no interaction with one another. Network virtualization is covered later in this chapter.

# Hypervisors

The software that makes server virtualization possible is called a *hypervisor*. The hypervisor acts as the abstraction layer between the physical hardware and the

guest operating system, which runs inside a virtual machine (VM). A VM is a logical container that typically has an operating system such as Windows Server or Linux with an application installed inside, similar to the installation of Windows Server or Linux and an application on physical hardware to serve client requests. The hypervisor is a thin operating system that sits between the physical hardware and the virtual machine and is meant to support only virtual machine operations and not any other applications. Alternatively, a hypervisor can be an application running inside a host operating system that allows for the creation of virtual machines. The placement and installation of a hypervisor within the virtualization infrastructure depends on the type of hypervisor.

There are two types of hypervisors: type 1 hypervisors (also known as bare-metal, or native, hypervisors) and type 2 hypervisors (also known as hosted hypervisors).

A type 1 hypervisor is a lightweight operating system that runs directly on the server hardware and provides the abstraction layer to create multiple virtual machines. The hypervisor component that conveys commands from the VM to the physical hardware that the hypervisor sits on is a virtual machine manager (VMM).

A type 1 hypervisor provides better scalability and performance than a type 2 hypervisor because it runs directly on top of the hardware and not inside another Windows or Linux operating system. The workload or VMs running on the hypervisor or host can scale across multiple physical servers, with hypervisors installed to provide high availability, clustering, and centralized manageability functions. A VM workload has full access to the underlying hardware resources. In addition, a type 1 hypervisor uses a small amount of memory and CPU overhead for its operations.

Examples of type 1 hypervisors include VMware ESXi, Microsoft Hyper-V, KVM, and Citrix Hypervisor (formerly Citrix XenServer).

A type 2, or hosted, hypervisor runs inside a Windows or Linux operating system installed on a physical host machine. It is known as a hosted hypervisor because the hypervisor is hosted within a host operating system. It relies on the host operating system to manage calls from VMs for CPU, memory, network, and storage resources. Type 2 hypervisors are usually deployed in smaller environments, mostly for testing and development purposes.

Examples of type 2 hypervisors are VMware Workstation, VMware Fusion, Parallels, and Oracle VirtualBox.

**ExamAlert**

Before taking the ENCOR exam, make sure you understand the differences between type 1 and type 2 hypervisors and their interaction with the underlying hardware resources.

Because type 1 hypervisors have direct access to host resources, they are usually more efficient than type 2 hypervisors. A type 1 hypervisor allows each guest operating system to interact directly with the hardware rather than going through a host operating system first; it is therefore more efficient. Type 2 hypervisor deployments require that the hypervisor application be installed onto a host operating system rather than directly on the hardware, as is done in a type 1 hypervisor deployment. With a type 2 deployment, the guest operating system does not directly access the host hardware due to that additional layer. Moreover, the host operating system requires a certain amount of CPU and memory resources to function optimally, and only the remaining amount is then available for virtual machine operations. When you deploy a traditional operating system with a type 2 hypervisor, this combination uses more hardware resources as overhead compared to running a thin type 1 hypervisor like VMware ESXi. Due to the higher overhead involved with type 2 hypervisors and the fact that the more advanced features of a virtualization platform are available only when running directly on hardware, the type 1 hypervisor is the more common type of hypervisor used in production virtualization deployments.

Figure 26.1 shows type 1 and type 2 hypervisors and their relationship to hardware and VMs. Note that with the type 1 hypervisor deployment, the hypervisor in running inside the physical server that is hosting the VMs. With the type 2 hypervisor, the hypervisor is running inside a host operating system, which is running inside the physical computer. The VMs are hosted within that hypervisor. Also, the type 2 hypervisor has one additional layer (the host operating system) than the type 1 hypervisor.

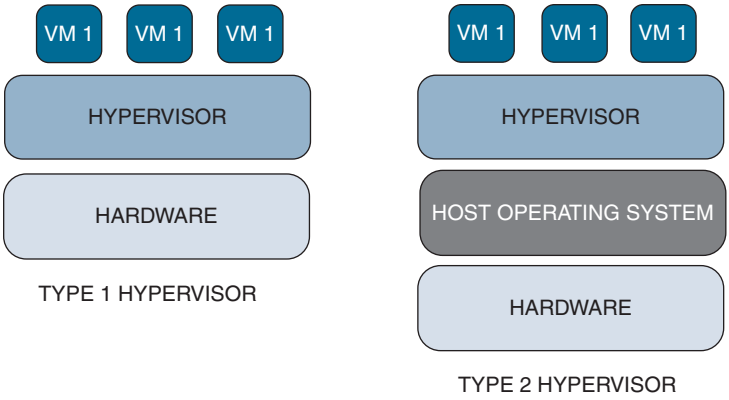



FIGURE 26.1 **Type 1 and Type 2 Hypervisors**

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which of the following is not an example of a type 1 hypervisor?
   - ❍ **A.** Citrix Hypervisor
   - ❍ **B.** VMware ESXi
   - ❍ **C.** Microsoft Hyper-V
   - ❍ **D.** VMware Fusion

2. Which of the following is a benefit of having a higher server consolidation ratio?
   - ❍ **A.** Higher ROI
   - ❍ **B.** Higher power consumption
   - ❍ **C.** Allows for mixing of different hypervisor types
   - ❍ **D.** Allows for the creation of larger VMs

3. Type 2 hypervisors are commonly deployed in what type of environment?
   - ❍ **A.** Production environment
   - ❍ **B.** Test/development environments
   - ❍ **C.** Service provider environments
   - ❍ **D.** Cloud service provider environments

# Answers

1. **D** is correct. VMware Fusion is a type 2 hypervisor that is built for use in a macOS environment
2. **A** is correct. A higher server consolidation ratio provides higher returns on capital expenditures.
3. **B** is correct. Type 2 hypervisors are typically deployed in non-production, or test/development, environments.

# Virtual Machines (VMs)

A virtual machine is a logical container that contains all the resources that an operating system requires for normal operation. Its components include virtual CPU (vCPU), virtual memory, a virtual graphic adapter, and a virtual NIC (vNIC), among other resources. An operating system and an application are installed into the VM. This operating system is termed the *guest operating system*, and it runs in a VM on a host.

Virtual machines are considered software representations of the CPU, memory, storage, and network resources that are used on a physical server. As far as the guest operating system is concerned, there is no difference between the virtual components it uses and physical hardware resources. In other words, the operating system and the application are not aware that they are running on virtualized resources rather than on physical server resources.

**ExamAlert**

Before taking the ENCOR exam, make sure you understand the high-level components that make up a virtual machine and how they are represented on the physical host/server.

The main components of a VM and how they are represented are highlighted in the following list:

- vCPUs, map to logical processors on the host CPU. A virtual machine can have one or more vCPUs assigned to it, depending on the CPU intensity of the application that is running within the VM.
- Virtual memory maps to memory on the physical host. However, memory is one resource that can be overprovisioned, which refers to allocating more memory to VMs than is installed on the physical host. Overprovisioning memory is acceptable provided that all the VMs do not use 100% of the memory assigned to them.
- A virtual disk is represented like any other disk to the guest operating system, but from the host's perspective, the virtual disk is represented as a file on the host-attached storage.
- A virtual NIC is a simulated virtual component on a host. A vNIC is attached to a virtual port on a virtual switch that sits within the hypervisor. Each virtual machine's vNIC has its own IP address and MAC address for communication purposes; this is similar to how a physical computer could have multiple NICs, each with a different IP address and a different

MAC address. The virtual switch has an uplink that maps to the physical NIC on the host. (This is also known as a VMNIC in a VMware ESXi environment.) Virtual switching is covered in the next section.

A VM can be considered the virtual equivalent of a physical server, and it requires the same software and network identifiers as a physical server. A VM uses traditional identifiers to communicate with other VMs or physical servers. However, one of the benefits of operating in a virtualized environment is that it gives you the flexibility to manipulate these values by basically editing the VM configuration file (for example, the MAC address of the VM). The same principle applies to the provisioning of new virtualized hardware for increasing or reducing a VM resource's capacity.

Management of the virtual machine environment is done in two ways. VMware ESXi is the most common hypervisor (type 1) in enterprise deployments, so it is covered here. VMware ESXi hosts can be managed directly using VMware Host Client, which is an HTML5-based client that can connect and manage a single ESXi host. VMware also supports vCenter Server, which is a centralized management server for managing multiple ESXi hosts and for the administration of more advanced services in the virtualization environment, such as vMotion (a vSphere feature that allows for live migration of running VMs from one host to another) and vSphere Distributed Switch (vDS). vCenter environment management is done either through the Flash-based vSphere Web Client or the HTML5-based vSphere HTML5 Web Client.

VMware previously had a vSphere C# client that needed to be installed on a Microsoft Windows system to manage an ESXi host or vCenter Server. You can still see this client being used in older VMware ESXi and vCenter Server deployments.

Like virtual machines, containers are isolated environments where applications run. Although containers have many similarities to virtual machines, they are not the same. Containers share the operating system and, where appropriate, bins/libraries.

A container typically contains an application and all the dependencies that the application needs to run. A container is physically smaller and more efficient than a VM, and it can be spun up more quickly, because it does not require deployment of an entire operating system and application. Platforms like Cisco Catalyst 3850 and 9300 running newer IOS XE software support a native lightweight Docker container that facilitates application-hosting environments like Guest Shell to run a Python environment. Guest Shell and Python are covered in Chapter 12, "Anatomy of Python."

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which of the following cannot be used to manage any VMware vCenter Server environment?
   - ❍ **A.** Flash-based vSphere Web Client
   - ❍ **B.** vSphere HTML5 Web Client
   - ❍ **C.** vSphere C# client
   - ❍ **D.** Microsoft Remote Desktop Protocol (RDP) client

2. What term applies to the operating system that is installed within a virtual machine?
   - ❍ **A.** Guest operating system
   - ❍ **B.** Host operating system
   - ❍ **C.** Thin operating system
   - ❍ **D.** Network operating system

3. True or false: Memory can be overprovisioned in a virtualized environment deployment.
   - ❍ **A.** True
   - ❍ **B.** False

## Answers

1. **D** is correct. The Microsoft RDP client cannot be used to manage a VMware vCenter Server environment.
2. **A** is correct. A guest operating system is installed inside a virtual machine.
3. **A** is correct. Memory can be overprovisioned. You can allocate more memory to VMs than you have installed on a physical host.

# Virtual Switching

VMware ESXi is the hypervisor typically found in enterprise virtualization environments, and this section covers virtual switching terminologies that are aligned with VMware. In a virtualization environment, rather than connecting directly to the physical Ethernet port of a virtual host, a VM vNIC connects to a virtual port on an internal switch called a *virtual switch*. A virtual switch, or vSwitch, as it is commonly termed, is a software-only Layer 2 switch that resides in a virtual host. Multiple vSwitches can exist in a virtual host, but they cannot share the same physical NIC. In addition, traffic cannot flow directly from one vSwitch to another. Due to this behavior of a vSwitch, there is no risk of a spanning tree loop within the virtual host.

A vSwitch connects the VMs running on a host to the outside world through use of a physical NIC in the virtual host as an uplink. When multiple uplinks are attached to a virtual host, the host can support teaming and redundancy for the vSwitch.

Virtual switches also support VLAN tagging. Ports with similar properties and segmentation requirements can be grouped together on a vSwitch. This grouping of ports is termed a port group in a VMware ESXi environment. When VMs need to communicate across port groups tagged in different VLANs, the traffic is treated like any other inter-VLAN traffic at an upstream Layer 3 device.

Figure 26.2 shows a VMware vSphere Standard Switch (vSS). It shows one virtual machine (vm1) connected to a VMware standard virtual switch (gray rectangle). The virtual switch has two NICs servicing it: vmnic4 and vmnic5. The three green indicators mean the VM, the vSwitch, and its NICs all currently have connectivity.

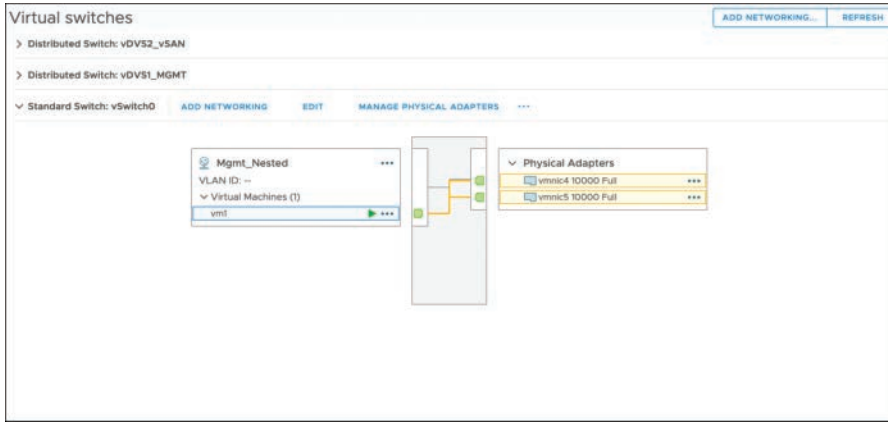



FIGURE 26.2 **vSphere Standard Switch**

So far, the discussion has been around vSphere Standard Switches, which are individual Layer 2 switches within an ESXi host. With a vCenter server, VMware supports one or multiple vSphere Distributed Switches (vDSs). A vDS provides a centralized interface to configure, administer, and monitor VM switching needs within the virtualization environment.

Figure 26.3 shows a vDS. As in the previous figure, one virtual machine is connected to a virtual switch. However, in this case, there is a distributed virtual switch that can service multiple virtualization hosts and their VMs.

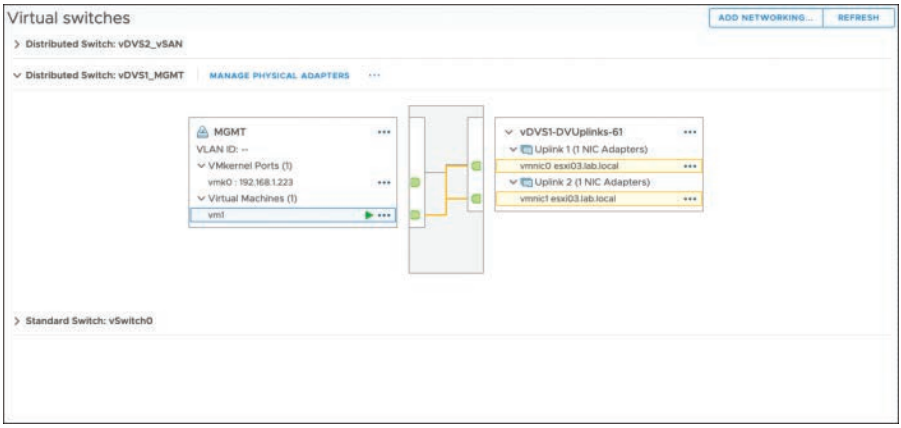



FIGURE 26.3 **vSphere Distributed Switch**

The following are examples of some common virtual switches:

- **Cisco ACI Virtual Edge:** This hypervisor-agnostic switch handles switching and policy enforcement for an ACI environment.
- **Cisco Nexus 1000VE:** This is a next-generation NX-OS CLI-based virtual switch for VMware vSphere environments. It is enhanced in the sense that it does not depend on in-kernel virtualization platform APIs, such as VMware APIs, as its predecessor Nexus 1000V (which is no longer supported by VMware) did.
- **vSphere Standard Switch (vSS):** vSS is used for simple vSphere deployments.
- **vSphere Distributed Switch (vDS):** vDS is used for complex vSphere environments with requirements for advanced networking and high availability.
- **Open vSwitch:** This multilayer switch is designed to operate in environments like XenServer, KVM, and Microsoft Hyper-V environments.

# Network Virtualization

Network virtualization is the formation of one or more logical isolated network overlays on top of a common physical network infrastructure (referred to as an *underlay* network). Each of the overlaid networks must provide the same services available on a traditional physical enterprise network infrastructure.

Network virtualization provides solutions to several business problems. One common issue that can be solved with network virtualization is secure visitor network access (via a guest network). Say that you need to give visitors Internet access while preventing unauthorized access to the internal enterprise network resources and services. This secure communication can be achieved by implementing a dedicated virtual network with a logical communication path for the entire guest network.

# Cisco Enterprise Network Function Virtualization (NFV)

Cisco Enterprise Network Function Virtualization (Enterprise NFV) addresses the requirements for deploying virtualized networks and application services, from orchestration and management to virtualization software. Enterprise NFV reduces operating and capital expenses and operational complexities for branch deployments by hosting network functions as software on a standard x86-based platform.

Network appliances such as firewalls and intrusion prevention systems (IPSs) are typically part of an enterprise branch architecture. In a high-availability environment, these devices are deployed in pairs for redundancy, further sprawling the hardware landscape in a branch site. Enterprise NFV can address some of these challenges.

These are some of the main features and benefits of Enterprise NFV:

- **Time savings:** Virtual network services can be deployed in minutes as opposed to the weeks or months that physical hardware deployments can take.
- **More choices:** An NFV infrastructure can be deployed on any general-purpose x86 platform. You can have a mix of virtual and physical devices, as well as both Cisco and third-party devices and network services.
- **Increased network uptime and deployment speed:** Enterprise NFV deployments are standardized through orchestration, thereby speeding maintenance and network incident response time. By using

a software-only approach to service delivery, network services can be deployed more quickly.

- **Reduction of operating expenses (opex) and capital expenses (capex):** With virtualization, costs are reduced and IT resources are freed up as a result of the reduction in hardware at the branches and in service visits for hardware installation and upgrades.
- **Enhanced network operations flexibility:** Flexibility increases thanks to the use of techniques such as virtual machine migration and snapshots.

# Cisco Enterprise NFV Architecture

Cisco Enterprise NFV is an architectural approach that seeks to decouple individual network services such as access control lists (ACLs), Network Address Translation (NAT), quality of service (QoS), intrusion prevention systems (IPSs), intrusion detection systems (IDSs), Layer 3 routing, wireless LAN controllers (WLCs), and more from the underlying hardware. Allowing these network functions to run inside virtual machines on a standard x86 platform increases deployment flexibility. NFV is typically accompanied by software-defined networking (SDN) approaches (that use orchestration and automation) to facilitate rapid deployment of network functions.

Cisco Enterprise NFV is an end-to-end solution that seeks to address all the requirements for deploying virtualized network and application services from management and orchestration to the virtualization package as well as options for the different hardware platforms.

The main building blocks of Cisco Enterprise NFV are as follows:

- **The orchestration environment:** This environment allows for the automation of the deployment of virtualized network services, consisting of different virtualized network functions.
- **Virtualized network functions (VNFs):** VNFs provide network functionality. ASAv and ISRv are examples.
- **NFV Infrastructure Software (NFVIS):** This platform supports the deployment and operations of the VNFs and hardware components.
- **x86 compute resources:** These resources provide the CPU, memory, and storage resources required to deploy and operate the various VNFs.

Figure 26.4 shows the building blocks of Cisco Enterprise NFV.

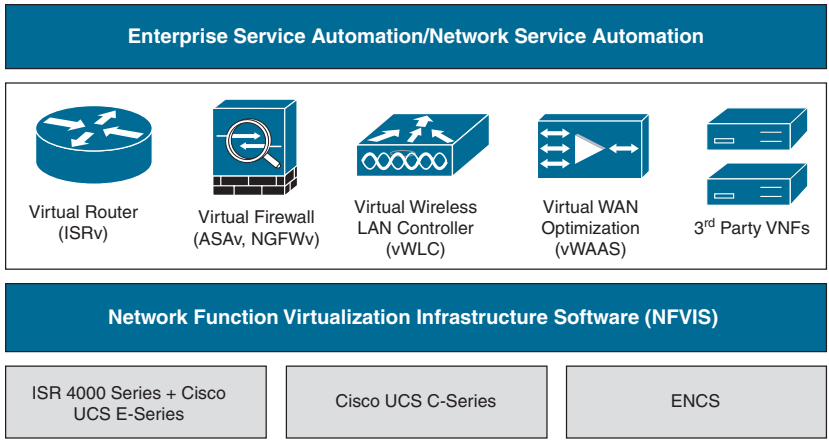



FIGURE 26.4 **Cisco Enterprise NFV Building Blocks**

# VNFs Supported in Cisco Enterprise NFV

Cisco Enterprise NFV supports virtualization for both network functions and applications in an enterprise branch. The following are some examples of VNFs that are supported in Cisco Enterprise NFV:

- Cisco Integrated Services Virtual Router (ISRv) for virtual routing
- Cisco Adaptive Security Virtual Appliance (ASAv), which is a firewall
- Cisco Firepower Next-Generation Firewall Virtual (NGFWv), which is an integrated virtual firewall with intrusion detection and intrusion prevention
- Cisco virtual Wide Area Application Services (vWAAS) for virtualized WAN optimization
- Cisco virtual Wireless LAN Controllers (vWLCs) for virtual WLCs

# Cisco NFV Hardware Options

The compute resources to operate a Cisco NFV are available as various platforms and form factors:

- **Cisco UCS C-Series servers:** These two- or four-socket rack servers can integrate with Cisco UCS fabric interconnects or can be deployed as standalone units.
- **Cisco ISR 4000 routers and Cisco E-Series servers:** When an integrated single-chassis solution is needed at a branch, the ISR 4000 is ideal because it can be used to virtualize network function using a Cisco UCS

E-Series server. ISR 4000 routers provide routing, hosting, security, and switching in a single platform. When equipped with NFVIS running on UCS E-Series modules, ISR 4000 routers provide optimally deployed network functions for branch sites.

- **Cisco ENCS:** Cisco 5100 and 5400 Enterprise Network Compute Systems (ENCS) are hybrid platforms that combine the best attributes of a traditional router and a traditional server and offer the same functionality with a smaller footprint. The NFVIS is used to facilitate the deployment and operation of VNFs and hardware components. ENCS x86 compute resources provide the CPU, memory, and storage resources required for the deployment and operation of these VNFs. ECNS can also accelerate some functions in hardware, such as inter-VM traffic flow, IP Security (IPsec) crypto, and RAID for storage.

Figure 26.5 shows the hardware options for Cisco NFV.

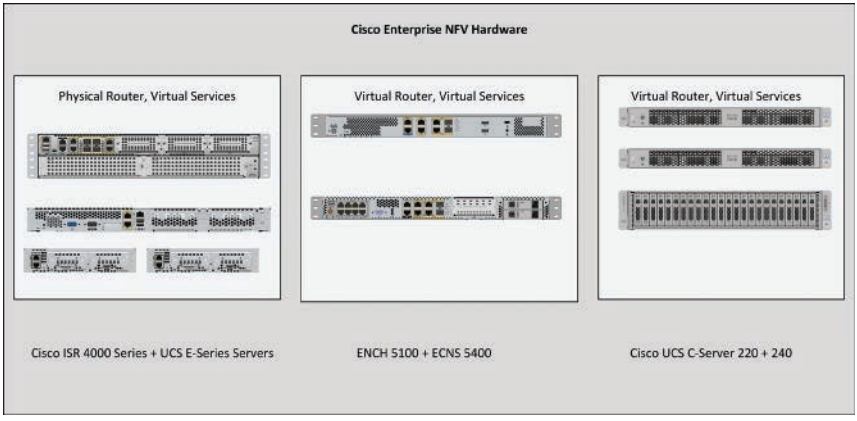



FIGURE 26.5 **NFV Hardware Options**

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. What is the term for the grouping of ports with similar properties within a VMware ESXi vSphere Standard Switch (vSS)?
   - ❍ **A.** vSphere Distributed Switch
   - ❍ **B.** VMkernel
   - ❍ **C.** Port group
   - ❍ **D.** Distributed port group

2. On which of the following virtual switches can the NX-OS CLI be used for administration?
   - ❍ **A.** Cisco Nexus 1000V
   - ❍ **B.** Open vSwitch
   - ❍ **C.** VMware VDS
   - ❍ **D.** Cisco Nexus 9000

3. Which Cisco virtual switch is used in a Cisco ACI environment?
   - ❍ **A.** Cisco ACI Virtual Edge
   - ❍ **B.** Cisco Nexus 1000VE
   - ❍ **C.** Open vSwitch
   - ❍ **D.** VMware VDS

4. A virtual switch in a VMware ESXi host has an uplink that maps to which of the following?
   - ❍ **A.** vNIC on the VM
   - ❍ **B.** Physical NIC or VMNIC
   - ❍ **C.** Virtual port on the virtual switch
   - ❍ **D.** None of the above

# Answers

1. **C** is correct. A port group is a grouping of ports with similar properties in a vSphere Standard Switch within a VMware ESXi host.
2. **A** is correct. The Cisco Nexus 1000V can be administered using the NX-OS CLI.
3. **A** is correct. Cisco Cisco ACI Virtual Edge can be deployed within a Cisco ACI environment.
4. **B** is correct. An uplink that services a virtual switch in the VMware ESXi environment is a physical NIC installed in that host, which is also known as a VMNIC.

# Review Questions

1. Which of the following is a logical container with virtualized resources that an operating system gets installed on to support client connections?
   - ❍ **A.** Host
   - ❍ **B.** Guest OS
   - ❍ **C.** VM
   - ❍ **D.** Hypervisor

2. VMware ESXi is an example of what type of hypervisor?
   - ❍ **A.** Type 1
   - ❍ **B.** Type 2
   - ❍ **C.** Hosted
   - ❍ **D.** VM

3. Which of the following are benefits of server virtualization? (Choose two.)
   - ❍ **A.** Hardware resource consolidation
   - ❍ **B.** Underutilization of CPU
   - ❍ **C.** Underutilization of memory
   - ❍ **D.** Reduction of TCO due to a less intense cooling requirement
   - ❍ **E.** Complete removal of physical networking

4. Which of the following is not a common virtual switch?
   - ❍ **A.** Cisco Nexus 1000V
   - ❍ **B.** Cisco Nexus 1000VE
   - ❍ **C.** Open vSwitch
   - ❍ **D.** VMware VDS
   - ❍ **E.** Cisco Nexus 9000

5. True or false: NFV infrastructure can be deployed on any general-purpose x86 platform.
   - ❍ **A.** True
   - ❍ **B.** False

# Answers to Review Questions

1. **C** is correct. A virtual machine is a logical container with virtualized CPU, memory, network, and disk resources and an operating system and application that serves clients.
2. **A** is correct. VMware ESXi is an example of a type 1 hypervisor. A type 1 hypervisor is a lightweight operating system that runs directly on server hardware and provides an abstraction layer for running virtual machines.
3. **A** and **D** are correct. Server virtualization consolidates multiple servers (virtual in this case) onto fewer powerful physical host; as a result of decreased physical server sprawl and reduced cabling, less cooling is needed in the data center, which reduces TCO.
4. **E** is correct. Cisco Nexus 9000 switches are scalable and high-density physical switches for data center deployments.
5. **A** is correct. NFV infrastructure can be deployed on any general-purpose x86 platform. There can be a mix of virtual and physical devices and Cisco and third-party devices and network services.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers VRF, GRE, and IPsec.

*This page intentionally left blank*

CHAPTER 27

# VRF Instances, GRE, and IPsec

**This chapter covers the following official ENCOR 350-401 exam objectives:**

- 2.2 Configure and verify data path virtualization technologies
- 2.2.a VRF
- 2.2.b GRE and IPsec tunneling

This chapter covers virtual pathing, which is a path isolation overlay network technique that is used for creating independent traffic paths to isolate traffic belonging to separate groups over a common enterprise physical network infrastructure. Several methods can be used to deploy path isolation in a campus network, including VRF-Lite and GRE tunnels. This chapter covers these methods along with IP Security (IPsec) VPNs.

**This chapter covers the following technology topics:**

- Virtual Routing and Forwarding (VRF)
- Generic Routing Encapsulation (GRE)
- IPsec VPNs

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. Which planes of operation can be virtualized?
2. What feature may need to be adjusted on a link to account for the additional overhead created by GRE?
3. What aspect of security ensures that data has not been altered or modified during transmission?

### Answers

1. Control plane and data plane
2. Maximum transmission unit (MTU)
3. Data integrity

# Virtual Routing and Forwarding (VRF)

Virtual routing and forwarding (VRF) involves creating multiple virtual networks within a single network entity. VRF instances are typically defined on a network device that serves as a boundary between the client-side VLAN and the Layer 3 network. Each VRF instance consists of an IP routing table, a forwarding table, and the interface or interfaces assigned to the instance. You can think of or compare a VRF instance to a virtual router residing within a Layer 3 switch or router. And as with any other Layer 3 devices, you can use routing protocols such as OSPF, EIGRP, and BGP for advertising the learned routes to populate the routing tables of these virtual networks.

You can compare the VRF instances on a router to virtual local area networks (VLANs) on a switch. However, instead of depending on Layer 2 technologies such as spanning tree, VRF instances allow for interaction and segmentation with Layer 3 dynamic routing protocols. The use of routing protocols over Layer 2 technologies has some advantages, such as improved network convergence times, dynamic traffic load sharing, and troubleshooting.

Figure 27.1 shows a comparison of a router without a VRF configuration and one with a VRF instance configured for management traffic.

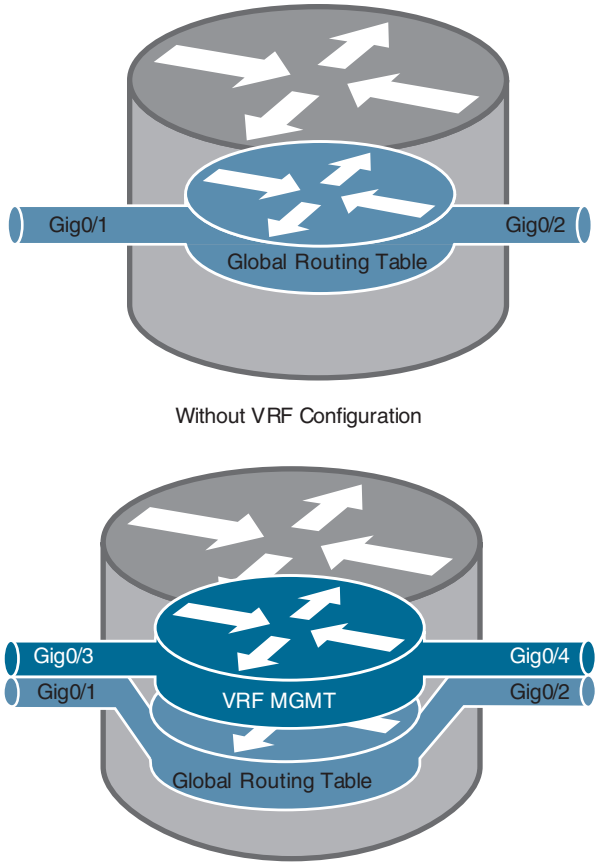



FIGURE 27.1 **Comparison of a Router with a VRF Instance and One Without VRF**

**ExamAlert**

Before taking the ENCOR exam, make sure you understand the use cases for VRF, the advantages they provide, and how they can be configured and verified.

# VRF-Lite

VRF-Lite uses a combination of VRF instances and 802.1Q trunking for hop-by-hop path isolation or Generic Routing Encapsulation (GRE)/Multipoint GRE (mGRE) for multi-hop path isolation. You typically find 802.1Q in an enterprise campus environment where the IP routing is under the control of

the organization deploying it, typically used with an IGP for routing. In a VRF-Lite end-to-end deployment, each Layer 3 device defines several VRF instances and leverages unique 802.1Q tags to send traffic belonging to separate VRF instances to the neighbor devices.

MPLS can also be used to extend VRF instances across an IP infrastructure. MPLS is typically used in large-scale environments where there are many virtual networks per device that need traffic engineering and optimal path selection using fast rerouting capabilities and the ability to provide full-mesh connectivity.

The use of VRF-Lite technology has the following advantages:

- **True routing and forwarding separation:** This means that dedicated control and data planes are used to handle traffic from different groups with various requirements. This segmentation provides an extra level of security because there is no communication between devices on different VRF instances unless explicitly allowed.
- **Simplified management and troubleshooting:** Management and troubleshooting of traffic are simplified because separate forwarding tables are used to switch traffic that belongs to a specific VRF instance; you are guaranteed that configuring overlay networks will not cause issues such as routing loops in the underlying global routing table.
- **Support for alternate default routes:** Due to the separate control and data planes, you can define separate default routes for each virtual network (that is, VRF instance).

In a network device, two planes of operation can be virtualized:

- **Control plane:** The control plane is a collection of processes that run at the process level and provide control function, including those related to making forwarding decisions and maintaining a network topology that is free of loops and unintended blackholes. All traffic that is directly or indirectly destined for a router is handled by the control plane. There is a requirement to virtualize the control plane because the virtual device must have a unique depiction of the virtual networks it handles.
- **Data plane:** The data plane, also known as the forwarding plane, is where the data actually flows. When a router receives a packet to route or a switch receives a frame to be switched, this occurs in the data plane. Information learned in the control plane facilitates data plane operation. Each virtual network has a unique forwarding table that needs to be virtualized.

The control plane and the data plane can be virtualized at different levels and map directly to different layers of the OSI model. For example, a VLAN-aware device can be virtualized at Layer 2 and yet have a single routing table; this means it is not virtualized at Layer 3. The various levels of virtualization may be useful depending on the technical requirements of a deployment. In some cases, Layer 2 virtualization may be enough, and in other instances, virtualization at different layers may be necessary. For example, to provide virtual firewall services, virtualization at Layers 2, 3, and 4 may be necessary; to define application services on each virtual firewall, Layer 7 virtualization may be required.

Let's look at an example of VRF-Lite configuration and verification. Example 27.1 shows how IP addresses are assigned in the global routing table as compared to within VRF instances. A VRF instance named MGMT is created, and two IP addresses are assigned to it and overlap with the ones configured in the global routing table. Because the MGMT VRF instance uses its own routing table, there is not a conflict between addresses. For verification, you can compare the global routing table with the MGMT VRF instance to see how the IP address and routes are shown. Notice the overlapping IP address ranges.

EXAMPLE 27.1 **Configuring IP Addresses in the Global Routing Table and Within a VRF Instance**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# interface GigabitEthernet0/0
R1(config-if)# ip address 10.10.10.1 255.255.255.0
R1(config-if)# interface GigabitEthernet0/1
R1(config-if)# ip address 10.10.20.1 255.255.255.0
R1(config-if)# vrf definition MGMT
R1(config-vrf)# address-family ipv4
R1(config-vrf-af)# interface GigabitEthernet0/2
R1(config-if)# vrf forwarding MGMT
R1(config-if)# ip address 10.10.10.1 255.255.255.0
R1(config-if)# interface GigabitEthernet0/3
R1(config-if)# vrf forwarding MGMT
R1(config-if)# ip address 10.10.20.1 255.255.255.0
R1(config-if)# end
R1#
R1# show ip route
Codes: L - local, C - connected, S - static, 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
       i - IS-IS, su - IS-IS summary, 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, H - NHRP, l
- LISP
       a - application route
       + - replicated route, % - next hop override, p - overrides from
PfR

Gateway of last resort is not set

      10.0.0.0/8 is variably subnetted, 4 subnets, 2 masks
C        10.10.10.0/24 is directly connected, GigabitEthernet0/0
L        10.10.10.1/32 is directly connected, GigabitEthernet0/0
C        10.10.20.0/24 is directly connected, GigabitEthernet0/1
L        10.10.20.1/32 is directly connected, GigabitEthernet0/1
R1#
R1# show ip route vrf MGMT

Routing Table: MGMT
Codes: L - local, C - connected, S - static, 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
       i - IS-IS, su - IS-IS summary, 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, H - NHRP, l
- LISP
       a - application route
       + - replicated route, % - next hop override, p - overrides from
PfR

Gateway of last resort is not set

      10.0.0.0/8 is variably subnetted, 4 subnets, 2 masks
C        10.10.10.0/24 is directly connected, GigabitEthernet0/2
L        10.10.10.1/32 is directly connected, GigabitEthernet0/2
C        10.10.20.0/24 is directly connected, GigabitEthernet0/3
L        10.10.20.1/32 is directly connected, GigabitEthernet0/3
R1#
```

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Each VRF instance contains all except which of the following components?
   - ❍ **A.** Routing table
   - ❍ **B.** Forwarding table
   - ❍ **C.** Interface assigned to the VRF instance
   - ❍ **D.** ACL

2. Which of the following planes of operation includes protocols, databases, and tables that are involved in making forwarding decisions and maintaining a functional network topology that is loop free?
   - ❍ **A.** Control plane
   - ❍ **B.** Management plane
   - ❍ **C.** Data plane
   - ❍ **D.** CEF

## Answers

1. **D** is correct. A VRF instance consists of an IP routing table, a derived forwarding table, and an interface or interfaces that use the forwarding table.
2. **A** is correct. The control plane includes all the protocols, databases, and tables involved in making forwarding decisions and maintaining a network topology that is free of loops or unintended blackholes.

# Generic Routing Encapsulation (GRE)

Generic Routing Encapsulation (GRE) is a tunneling protocol that provides a path for transporting packets over an IP network by encapsulating a packet inside a transport protocol. GRE encapsulates a payload—that is, an inner packet that needs to be delivered to a destination network inside an outer IP packet. The GRE tunnel behaves as a virtual point-to-point link with two endpoints identified by a tunnel source and a tunnel destination address. The tunnel endpoint sends payload through the tunnel by routing encapsulated packets through intervening IP networks. Routers along this path do not parse the payload but only parse the outer IP packet to forward it toward the GRE tunnel endpoint. When the IP packet arrives at the tunnel destination, the GRE encapsulation is removed, and the payload is forwarded to the packet's ultimate destination. The encapsulation point is called the *tunnel entry*, and the decapsulation point is called the *tunnel exit*.

Figure 27.2 shows a topology in which R1 and R2 connect to their respective ISPs using their ISPs as their default gateway. This allow R1 and R2 to reach each other's Internet-facing interfaces (Gi0/0 on both routers) to form a GRE tunnel over the Internet. Network 10.10.10.0/24 is the underlay network, and network 100.1.1.0/24 is the GRE tunnel (that is, overlay network).

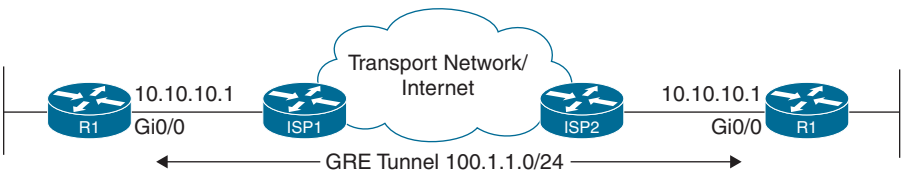



FIGURE 27.2 **GRE Tunnel Topology**

Some of the key benefits of encapsulating traffic in another network layer protocol are as follows:

- It provides multiprotocol local networks over a single-protocol backbone.
- It provides a communication path for networks that contain protocols with limited hop count. For example, if the two hosts' communication paths cannot exceed 15 hosts, a tunnel can be created to hide some of the hops inside the network.
- It allows VPNs across WANs.
- It can connect distant subnetworks.

The tunneled packets logically consist of the following:

- **Payload data:** This is the data that will be tunneled.
- **Encapsulation header:** This specifies additional control information about the payload that is being carried or the forwarding behavior that needs to be applied to the packets being tunneled at decapsulation—or both.
- **Delivery or transport header:** This indicates how the encapsulated payload data is transported to the other end of the tunnel.

At a high level, GRE is primarily intended to allow a device running a network protocol to communicate over a network running a different network layer protocol. GRE tunnels have some limitations and key characteristics.

ExamAlert

Before taking the ENCOR exam, make sure you understand the key characteristics and limitations of GRE tunnels.

The following are some of the key characteristics and limitations of GRE tunnels:

- **Stateless tunnels:** The tunnel endpoint does not keep information about the remote tunnel endpoint's state or availability.
- **At least 24 bytes of overhead:** This includes a new 20-byte IP header, which indicates the source and destination IP addresses of the GRE tunnel. The remaining 4 bytes are for the GRE header itself.
- **Multiprotocol and ability to tunnel any OSI Layer 3 protocol:** It uses a protocol type field in the GRE header to support the encapsulation of any OSI Layer 3 protocol.
- **Routing protocols carried within the tunnel:** GRE allows routing protocols (such as OSPF and EIGRP) to be tunneled across the connection.
- **Weak security features:** No strong confidentiality, data source authentication, or data integrity mechanism exists in GRE. GRE can, however, be used with IPsec to provide confidentiality, source authentication, and data integrity.
- **Protocol:** GRE packets that are encapsulated within IP use IP protocol type 47.

With GRE, the larger packet size created by the additional headers can cause a network performance hit. If the packets are larger than the interface's maximum transmission unit (MTU) permits, the router must fragment the packet. This fragmentation effort can cause significant CPU overhead on a router, which can affect all packet forwarding. Therefore, you may need to adjust the MTU on the GRE tunnel, and the MTU needs to match at both ends.

You change the MTU on an interface by using the command **ip mtu** *mtu*. Sometimes ICMP messages may be blocked and must be relayed from the host for the path MTU to work. In such a case, you also need to adjust the TCP maximum segment size (MSS) value to prevent TCP sessions from being dropped; to do so, you adjust the MSS value of the TCP SYN packets by using the command **ip tcp adjust-mss** *mss value* for the interface.

Figure 27.3 shows a GRE packet.

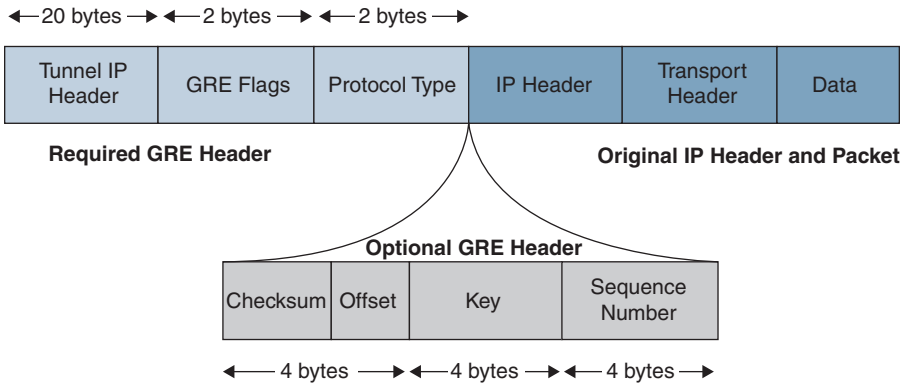



FIGURE 27.3 **GRE Packet Format**

Example 27.2 shows the configuration of a basic GRE tunnel, using the topology shown in Figure 27.2. Interface Tunnel 0 is created and assigned an IP address in the 100.1.1.0/24 network, which would be considered the overlay network. IP addresses in the 10.10.10.0/24 would be the source of the traffic (that is, the underlay network). The tunnel destination is the IP address at the other end of the underlay network for both R1 and R2.

EXAMPLE 27.2 **Configuring a Basic GRE Tunnel**

```
R1#
R1# config terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# interface tunnel 0
R1(config-if)# ip address 100.1.1.1 255.255.255.0
```

```
R1(config-if)# tunnel source 10.10.10.1
R1(config-if)# tunnel destination 10.10.10.2
R1(config-if)# end
R1#

R2#
R2# config terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R2(config)# interface tunnel 0
R2(config-if)# ip address 100.1.1.2 255.255.255.0
R2(config-if)# tunnel source 10.10.10.2
R2(config-if)# tunnel destination 10.10.10.1
R2(config-if)# end
R2#
```

Example 27.3 shows the configuration of a basic GRE tunnel in a VRF environment. This example, which builds on the VRF and GRE examples already shown in this chapter, involves creating a VRF instance (GREEN) and then using GRE to have complete routing isolation.

EXAMPLE 27.3 **Configuring a Basic GRE Tunnel in a VRF Environment**

```
R1#
R1# config terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# vrf definition GREEN
R1(config-vrf)# address-family ipv4
R1(config-vrf-af)# exit-address-family
R1(config-vrf)# exit
R1(config)# interface tunnel 0
R1(config-if)# vrf forwarding GREEN
R1(config-if)# ip address 100.1.1.1 255.255.255.0
R1(config-if)# tunnel source 10.10.10.1
R1(config-if)# tunnel destination 10.10.10.2
R1(config-if)# end
R1#

R2#
R2# config terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R2(config)# vrf definition GREEN
R2(config-vrf)# address-family ipv4
R2(config-vrf-af)# exit-address-family
R2(config-vrf)# exit
R2(config)# interface tunnel 0
R2(config-if)# vrf forwarding GREEN
R2(config-if)# ip address 100.1.1.2 255.255.255.0
```

```
R2(config-if)# tunnel source 10.10.10.2
R2(config-if)# tunnel destination 10.10.10.1
R2(config-if)# end
R2#
```

Let's now look at a couple commands for troubleshooting GRE tunnels. Some common issues you may encounter include the following:

- GRE source IP address not reachable by the remote host
- GRE destination IP address not reachable by the local host
- Recursive routing

To quickly get to the bottom of some common GRE tunnel problems, you can use the following commands:

- **debug tunnel:** This command allows you to get debugging information and events related to the tunnel quickly.
- **debug tunnel packet:** This command enables you to quickly get packet debugging information and events related to the tunnel packets.

Example 27.4 shows the verification of a GRE tunnel, using the **show interface tunnel** *number* command. The output of this command gives you a view of the interface status, IP address, tunnel mode, and tunnel source and destination.

EXAMPLE 27.4 **Verifying a GRE Tunnel**

```
R1#
R1# show interfaces tunnel 0
Tunnel0 is up, line protocol is up
  Hardware is Tunnel
  Internet address is 100.1.1.1/24
  MTU 17916 bytes, BW 100 Kbit/sec, DLY 50000 usec,
     reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation TUNNEL, loopback not set
  Keepalive not set
  Tunnel linestate evaluation up
  Tunnel source 10.10.10.1, destination 10.10.10.2
  Tunnel protocol/transport GRE/IP
    Key disabled, sequencing disabled
    Checksumming of packets disabled
  Tunnel TTL 255, Fast tunneling enabled
  Tunnel transport MTU 1476 bytes
<...output omitted...>
```

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which of the following is not a characteristic of GRE?
   - ❍ **A.** GRE has weak security features.
   - ❍ **B.** GRE can tunnel any Layer 3 protocol.
   - ❍ **C.** GRE tunnels are stateful.
   - ❍ **D.** GRE can carry routing protocols within the tunnel.

2. Which of the following is not necessary to create a GRE tunnel?
   - ❍ **A.** IP address
   - ❍ **B.** Tunnel source
   - ❍ **C.** Tunnel destination
   - ❍ **D.** Routing protocol

3. Which of the following is performed by the routers in a tunnel path between two GRE router endpoints?
   - ❍ **A.** Parsing the payload
   - ❍ **B.** Parsing the IP packet
   - ❍ **C.** Encapsulating the payload packet
   - ❍ **D.** Decapsulating the payload packet

## Answers

1. **C** is correct. GRE tunnels are stateless, meaning the tunnel endpoint does not keep any information about the state or availability of the remote tunnel endpoint.
2. **D** is correct. A routing protocol is not necessary for the creation of GRE tunnels.
3. **B** is correct. Routers along a tunnel path only parse the outer IP packet as they forward it toward the GRE tunnel endpoint.

# IPsec VPNs

**ExamAlert**

Before taking the ENCOR exam, make sure you know the various VPN deployment types for Cisco IOS.

This section covers site-to-site VPNs, virtual tunnel interfaces (VTIs) on Cisco IOS devices, Dynamic Multipoint VPN (DMVPN), FlexVPN and IPsec, and GRE tunneling over IPsec.

## Site-to-Site VPNs

Enterprises typically used site-to-site VPNs to replace traditional WAN networks to connect to their geographically dispersed sites or partner networks over the Internet or over an MPLS WAN. A site-to-site VPN provides the benefit of reduced cost as sites are brought online, especially if they are established over the Internet.

One common deployment scenario for site-to-site VPNs is a headquarters network providing remote offices access to the corporate intranet. In this scenario, the headquarters and remote office may be connected through a GRE tunnel established over the Internet or MPLS WAN. The remote office branch users can access all network services at the headquarters. You typically see a scenario like this using GRE tunneling over IPsec (covered later in this chapter).

VPN topologies are logical connections between networks, and thus, the physical network or interconnection of network devices has no bearing on how the VPN protocols create connections between networks. The following VPN topologies are commonly used in site-to-site VPNs (see Figure 27.4):

- **An individual point-to-point network:** Multiple sites can be interconnected using one or more point-to-point VPN connections.
- **A hub-and-spokes network:** One site act as the hub, and all other peer sites act as spokes. Traffic flow can be from spoke to hub and then to spoke, or the hub can act as a relay to facilitate spoke-to-spoke communication.
- **A fully meshed network:** Every device has a connection to every other device. This essentially means that each site can communicate with every other site in the same VPN.

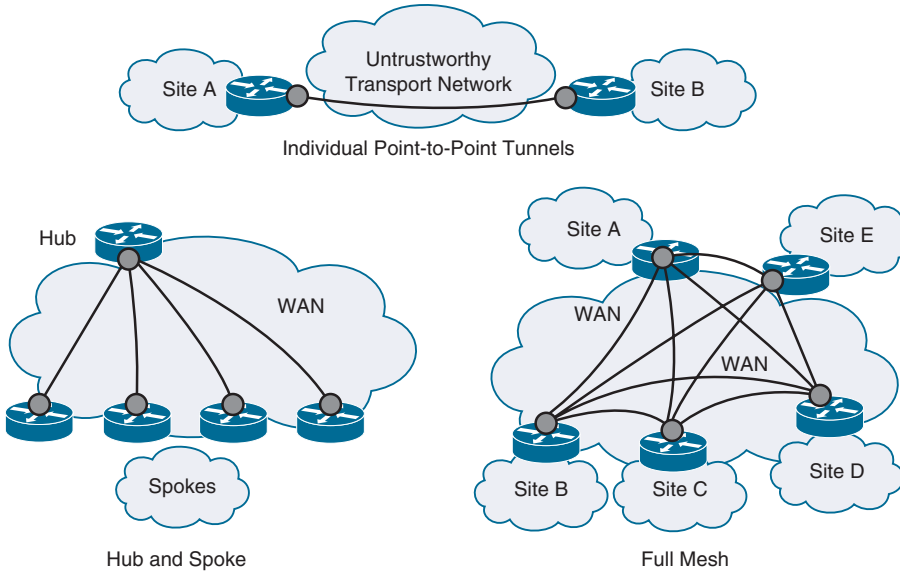



FIGURE 27.4 **Logical VPN Topologies Used in Site-to-Site VPNs**

# Dynamic Multipoint VPN (DMVPN)

Dynamic Multipoint VPN (DMVPN) is a centralized architecture that provides ease of management and control for VPN deployments. It allows for easier scaling of large and small IPsec VPNs with GRE tunnels, IPsec encryption, and Next Hop Resolution Protocol (NHRP). DMVPN is ideally used in organizations that require WAN connectivity between remote sites. Considerations for its use include cost-driven use of the Internet to replace a dedicated private WAN link or achieving regulatory compliance where encryption of private WAN links is required.

Some of the high-level benefits of using DMVPN are as follows:

- Zero-touch provisioning for the addition of new remote sites
- Automatic IPsec triggering for the building of IPsec tunnels between sites
- On-demand full-mesh connectivity with a simple hub-and-spoke configuration
- Reduced latency and bandwidth savings
- Reduced deployment complexity due to more straightforward configurations
- Improved resiliency, thanks to the dual-hub configuration

Figure 27.5 shows a DMVPN scenario with dynamic site-to-site tunnels established from spoke to hub and spoke to spoke, as needed.

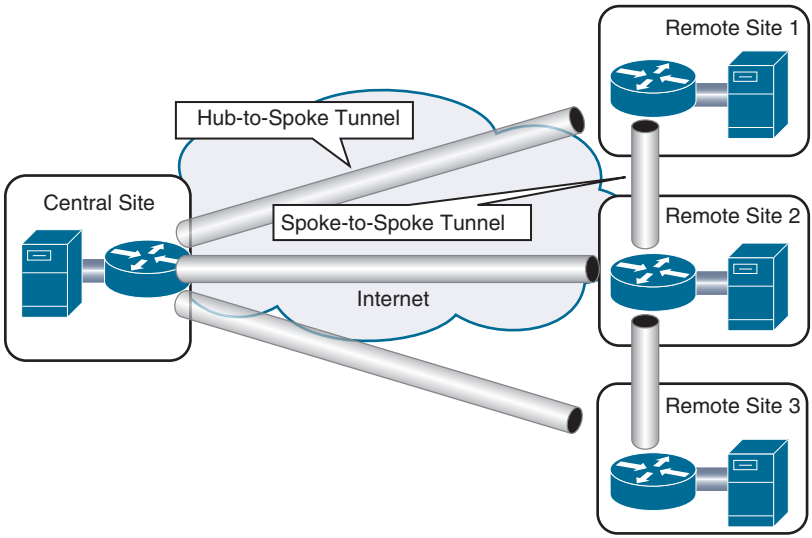



FIGURE 27.5 **Cisco DMVPN Topology**

# Cisco IOS Virtual Tunnel Interfaces (VTIs)

Cisco IOS VTIs are used to provide a simplified configuration process for connecting remote sites. VTIs provide an alternative to the use of GRE or Layer 2 Tunneling Protocol (L2TP) for encapsulation. An IPsec VTI provides the flexibility of forwarding both IP unicast and IP multicast encrypted traffic on any physical interface with multiple paths. The traffic is encrypted and decrypted when forwarded to and from the tunnel interface and is managed by the IP routing table. Using the IP routing table to forward traffic to the tunnel interface simplifies IPsec VPN configuration compared to using crypto maps and access control lists in traditional IPsec configurations.

The following are some of the main characteristics and benefits of using VTIs:

- **Simplified management:** The simplified construct of a Cisco IOS VTI translates to reduced costs and downtime due to the simplification of the configuration compared to the traditional cryptographic map style of configuration. The configuration on the IOS CLI is more intuitive.
- **Support for multicast encryption:** Customers can use VTIs to securely transport multicast traffic, control traffic, or data traffic from one site to another.

- **Routable interfaces:** A Cisco IOS IPsec VTI can support all types of IP routing protocols.
- **Improved scaling:** Cisco IOS VTIs need fewer established IPsec security associations to cover different types of traffic (both unicast and multicast), thus enabling improved scaling.

Figure 27.6 shows a comparison of GRE packet encapsulation and IPsec tunnel mode with a VTI.

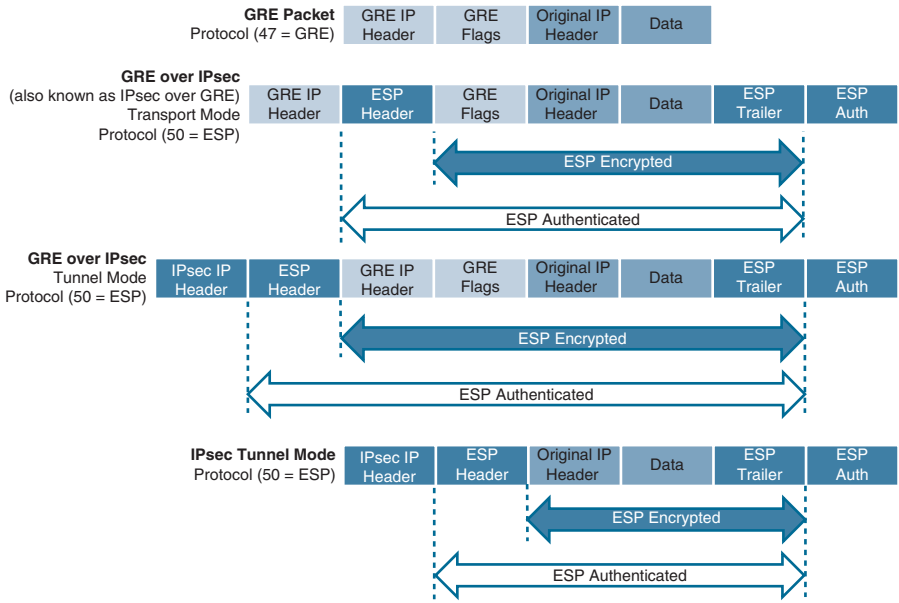



FIGURE 27.6 **Comparison of GRE Packet Encapsulation and IPsec Tunnel Mode**

# Cisco IOS FlexVPN

Cisco IOS FlexVPN is a unified solution that simplifies the deployment of multiple types of VPNs. FlexVPN is a unified ecosystem that addresses the complexities of deploying multiple different types of VPNs. It encompasses all types of VPNs, including site-to-site and remote access VPNs.

FlexVPN uses standard-based encryption technology to allow larger organizations to connect branch offices and remote users securely. It introduces significant cost savings compared to managing multiple VPN solutions, such as GRE, crypto maps, and VTI-based solutions. Cisco FlexVPN uses open

standards–based IKEv2 for security, and this provides compatibility with any IKEv2-based third-party VPN vendors, including a multitude of VPN clients.

The following are the main features and benefits of Cisco FlexVPN:

- **Transport network:** FlexVPN can be deployed over the public Internet or a private MPLS network.
- **Deployment flexibility:** A single FlexVPN deployment can be deployed to accept both site-to-site and remote access VPNs at the same time.
- **Failover redundancy:** Several redundancy models can be deployed with FlexVPN: dynamic routing protocols over the FlexVPN tunnel where the path selection is based on the routing metrics, IKEv2-based route distribution and server cluster, and IKEv2 active/standby stateful failover between two chassis.
- **Multicast support:** FlexVPN natively supports IP multicast.
- **Quality of service (QoS):** FlexVPN allows for hierarchical QoS to be integrated per tunnel or SA.
- **VRF awareness:** FlexVPN can integrate with MPLS VPN for support in service provider–type deployments.
- **Centralized policy control:** VPN dynamic policies such as VRF selection, DNS servers (for remote access), and split-tunnel policy can be fully integrated with AAA servers and applied on a per-peer basis.

# IP Security (IPsec)

IPsec is a suite of protocols based on RFC 4301 that provides data confidentiality, data integrity, and data origin authentication of IP packets. Corporate data that is sensitive must remain private when it is transmitted over a public network. IPsec provides security at the IP layer and offers a standards-based approach to providing secure and private data transmission. IPsec also introduces cost savings compared to traditional WAN access methods since most of its implementation is over the Internet.

Typically, IPsec is associated with virtual private networks (VPNs), as it basically creates private connections or networks between two endpoints. IPsec includes the following features:

- **Data confidentiality:** IPsec keeps data within an IPsec VPN private between the participants of that VPN. Most VPNs are used across the public Internet, and the data may be intercepted and examined, so there is a need for data confidentiality.

- **Data integrity:** IPsec guarantees that the data was not altered or modified during transmission through an IPsec VPN. Data integrity is typically verified using a hash algorithm to check that the data was not modified between endpoints. Packets that are found to have been changed during transmission are dropped.
- **Data origin authentication:** IPsec validates the source of a transmission. Each end of the VPN does this validation to make sure that the other end is what it wants to be connected to.
- **Anti-replay:** IPsec ensures that no packets within a VPN are duplicates. This protection is based on the use of sequence numbers in the packets and a sliding window at the receiver side. The sequence number is compared to the sliding window to determine packets that are late. Late packets are dropped because they are considered duplicates.

The features of IPsec are implemented in a series of standards-based protocols. The IPsec protocols do not specify any authentication, encryption algorithms, key generation techniques, or security association (SA) mechanisms.

IPsec uses three main protocols—IKE, AH, and ESP—which are covered next.

## Internet Key Exchange (IKE) Version 1 (IKEv1) and Version 2 (IKEv2)

IKE uses UDP port 500 and provides authentication of IPsec peers, negotiates IPsec security associations, and establishes IPsec keys. It is a hybrid protocol that implements Oakley and SKEME key exchanges inside the Internet Security Association and Key Management Protocol (ISAKMP) framework. Although IKE can be used with other protocols, it was originally implemented with IPsec.

An IPsec VPN gateway uses either IKEv1 or IKEv2 to negotiate the IKE SA and IPsec tunnel. Some of the significant advantages of IKEv2 over IKEv1 are highlighted in the following list:

- **Support for NAT traversal (NAT-T):** IKEv2 introduces support for NAT-T, which is necessary when routers along the data path need to perform Network Address Translation.
- **Built-in health check:** IKEv2 can automatically reestablish a tunnel if it goes down by using a "liveness" check (which replaces the dead peer detection in IKEv1).
- **EAP authentication:** IKEv2 specifies that EAP must be used with public-key signature-based responder authentication.

## Authentication Header (AH)

AH uses protocol number 51 and provides encapsulation for user traffic authentication. It provides data integrity, data origin authentication, and protection against replay to user traffic. AH is not recommended for VPNs over untrusted networks because it does not provide encryption.

## Encapsulating Security Payload (ESP)

ESP uses protocol number 50 and provides encapsulation for user traffic encryption and authentication. ESP includes data integrity, data origin authentication, protection against replay, and confidentiality to user traffic. ESP is preferred over AH for VPNs over untrusted networks because it allows for data confidentiality through encryption.

## IPsec VPN Examples

Example 27.5 shows a basic site-to-site IPsec VPN configuration. This example configures a site-to-site IPsec tunnel using GRE over IPsec with a pre-shared key. R1 and R2 are configured for IPsec over GRE using crypto maps.

EXAMPLE 27.5 **Configuring a Basic Site-to-Site IPsec VPN**

```
R1#
R1# config terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# crypto isakmp policy 10
R1(config-isakmp)# encryption aes
R1(config-isakmp)# hash sha
R1(config-isakmp)# authentication pre-share
R1(config-isakmp)# group 2
R1(config-isakmp)# exit
R1(config)# crypto isakmp key 0 Cisco123 address 100.1.1.2
R1(config)# crypto ipsec transform-set TRANSFORMSET esp-aes
esp-sha-hmac
R1(cfg-crypto-trans)# crypto map CRYPTOMAP 10 ipsec-isakmp
% NOTE: This new crypto map will remain disabled until a peer
        and a valid access list have been configured.
R1(config-crypto-map)# set peer 100.1.1.2
R1(config-crypto-map)# set transform-set TRANSFORMSET
R1(config-crypto-map)# match address 100
R1(config-crypto-map)# exit
R1(config)# access-list 100 permit ip host 1.1.1.1 host 2.2.2.2
R1(config)# interface g0/0
R1(config-if)# crypto map CRYPTOMAP
```

```
R1(config-if)# exit
R1(config)# ip route 2.2.2.2 255.255.255.255 100.1.1.2
R1(config)# end
R1#


R2#
R2# config terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R2(config)# crypto isakmp policy 10
R2(config-isakmp)# encryption aes
R2(config-isakmp)# hash sha
R2(config-isakmp)# authentication pre-share
R2(config-isakmp)# group 2
R2(config-isakmp)# exit
R2(config)# crypto isakmp key 0 Cisco123 address 100.1.1.1
R2(config)# crypto ipsec transform-set TRANSFORMSET esp-aes
esp-sha-hmac
R2(cfg-crypto-trans)# crypto map CRYPTOMAP 10 ipsec-isakmp
% NOTE: This new crypto map will remain disabled until a peer
        and a valid access list have been configured.
R2(config-crypto-map)# set peer 100.1.1.1
R2(config-crypto-map)# set transform-set TRANSFORMSET
R2(config-crypto-map)# match address 100
R2(config-crypto-map)# exit
R2(config)# access-list 100 permit ip host 2.2.2.2 host 1.1.1.1
R2(config)# interface g0/0
R2(config-if)# crypto map CRYPTOMAP
R2(config-if)# exit
R2(config)# ip route 1.1.1.1 255.255.255.255 100.1.1.1
R2(config)# end
R2#
```

Example 27.6 shows the verification of the site-to-site IPsec configuration. As you can see, there is reachability over the tunnel. The output of the command **show crypto isakmp sa** shows the ISAKMP security association (SA) status as active and in a QM_IDLE state. The QM_IDLE state means that the SA remains authenticated with its peer and may be used for subsequent quick mode exchanges for more IPsec SAs. The output of **show crypto ipsec sa** displays information about the IPsec SA.

EXAMPLE 27.6 **Verifying a Site-to-Site IPsec Configuration**

```
R1#
R1# ping 2.2.2.2 source loopback0
Type escape sequence to abort.
```

```
Sending 5, 100-byte ICMP Echos to 2.2.2.2, timeout is 2 seconds:
Packet sent with a source address of 1.1.1.1
.!!!!
Success rate is 80 percent (4/5), round-trip min/avg/max = 6/7/10 ms
R1#
R1# show crypto isakmp sa
IPv4 Crypto ISAKMP SA
dst             src             state          conn-id status
100.1.1.2       100.1.1.1       QM_IDLE           1001 ACTIVE

IPv6 Crypto ISAKMP SA


R1# show crypto ipsec sa

interface: GigabitEthernet0/0
    Crypto map tag: CRYPTOMAP, local addr 100.1.1.1

   protected vrf: (none)
   local  ident (addr/mask/prot/port): (1.1.1.1/255.255.255.255/0/0)
   remote ident (addr/mask/prot/port): (2.2.2.2/255.255.255.255/0/0)
   current_peer 100.1.1.2 port 500
     PERMIT, flags={origin_is_acl,}
    # pkts encaps: 4, # pkts encrypt: 4, # pkts digest: 4
    # pkts decaps: 4, # pkts decrypt: 4, # pkts verify: 4
    # pkts compressed: 0, # pkts decompressed: 0
    # pkts not compressed: 0, # pkts compr. failed: 0
    # pkts not decompressed: 0, # pkts decompress failed: 0
    # send errors 0, # recv errors 0

     local crypto endpt.: 100.1.1.1, remote crypto endpt.: 100.1.1.2
     plaintext mtu 1438, path mtu 1500, ip mtu 1500, ip mtu idb
GigabitEthernet0/0
     current outbound spi: 0x5CA35446(1554207814)
     PFS (Y/N): N, DH group: none

     inbound esp sas:
      spi: 0x7EF923B3(2130256819)
        transform: esp-aes esp-sha-hmac ,
        in use settings ={Tunnel, }
        conn id: 1, flow_id: SW:1, sibling_flags 80004040, crypto map:
CRYPTOMAP
        sa timing: remaining key lifetime (k/sec): (4288244/3485)
        IV size: 16 bytes
        replay detection support: Y
        Status: ACTIVE(ACTIVE)

     inbound ah sas:
```

```
     inbound pcp sas:

     outbound esp sas:
      spi: 0x5CA35446(1554207814)
        transform: esp-aes esp-sha-hmac ,
        in use settings ={Tunnel, }
        conn id: 2, flow_id: SW:2, sibling_flags 80004040, crypto map:
CRYPTOMAP
        sa timing: remaining key lifetime (k/sec): (4288244/3485)
        IV size: 16 bytes
        replay detection support: Y
        Status: ACTIVE(ACTIVE)

     outbound ah sas:

     outbound pcp sas:
R1#
```

# GRE Tunneling over IPsec

GRE is typically used to tunnel unicast and multicast traffic between routers and for routing protocols between sites. One of the significant downsides of using GRE tunnels alone is that they transport packets in plaintext and offer no protection for the payload. You can use IPsec to secure the tunnel as you move data between tunnel endpoints by encrypting the entire GRE tunnel.

Although IPsec provides a secure method for tunneling data across an IP network, it has certain limitations. By itself, IPsec does not support IP broadcast or IP multicast, so it cannot be used with certain protocols, such as routing protocols. As you saw earlier in this chapter, GRE can carry other protocols, such as IP broadcast and IP multicast. Using GRE in conjunction with IPsec provides the ability to securely transport routing protocols, IP multicast, or multiprotocol traffic between a main site and branch sites.

With point-to-point GRE over IPsec, all traffic between sites is encapsulated in a point-to-point GRE packet before the encryption process, simplifying the access control list used in the crypto map statement. The crypto map needs only one line that permits GRE (IP protocol 47).

There are two IPsec modes:

- **Tunnel mode:** This mode protects the original IP header within a new IPsec IP header. Tunnel mode is the default mode.
- **Transport mode:** This mode is used if the original IP header can be exposed. Transport mode is normally sufficient with GRE over IPsec because the GRE and IPsec endpoints are often the same.

Figure 27.7 shows a GRE over IPsec packet. The innermost layer is the original IP packet, and it represents the data traveling between the two sites. The original IP packet is encapsulated in a GRE header to permit routing protocol transmission in the GRE tunnel. IPsec is then added as the outer layer to provide data confidentiality and integrity.

**Tunnel Mode**


**Transport Mode**


FIGURE 27.7 **GRE Over IPsec Packet Format**

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which of the following VPN technologies is used for the dynamic creation of IPsec tunnels between sites without the need for a permanent connection?
   - ❍ **A.** VTI
   - ❍ **B.** Cisco FlexVPN
   - ❍ **C.** DMVPN
   - ❍ **D.** Remote access VPN

2. Which VPN technology serves as a unifier for multiple types of VPNs?
   - ❍ **A.** VTI
   - ❍ **B.** Cisco FlexVPN
   - ❍ **C.** Site-to-site
   - ❍ **D.** Remote access VPN

3. Which key management protocol introduced support for NAT traversal?
   - ❍ **A.** AH
   - ❍ **B.** ESP
   - ❍ **C.** IKEv1
   - ❍ **D.** IKEv2

# Answers

1. **C** is correct. DMVPN allows for the automatic triggering of the creation of IPsec tunnels between sites.
2. **B** is correct. Cisco FlexVPN is a unified solution that simplifies the deployment of multiple types of VPNs.
3. **D** is correct. NAT-T was incorporated into IKEv2.

# Review Questions

1. True or false: VRF-Lite technology provides true routing and forwarding separation.
   - ❍ **A.** True
   - ❍ **B.** False

2. Which of the following is a multi-hop technique for data path virtualization?
   - ❍ **A.** mGRE
   - ❍ **B.** BGP
   - ❍ **C.** 802.1Q
   - ❍ **D.** Forwarding table

3. Which of the following is a characteristic of GRE?
   - ❍ **A.** GRE uses IP protocol 43
   - ❍ **B.** GRE uses IP protocol 47
   - ❍ **C.** GRE tunnels are stateful
   - ❍ **D.** GRE provides strong security features

4. Which of the following is not one of the protocols used with IPsec?
   - ❍ **A.** AH
   - ❍ **B.** ESP
   - ❍ **C.** IKE
   - ❍ **D.** EAP

# Answers to Review Questions

1. **A** is correct. VRF-Lite offers true routing and forwarding separation. It has dedicated control and data planes that are used to handle traffic from different groups with various requirements.
2. **A** is correct. mGRE is used for multi-hop path isolation to extend virtual networks.
3. **B** is correct. IP protocol 47 defines GRE packets.
4. **D** is correct. Authentication Header (AH), Encapsulating Security Payload (ESP), and Internet Key Exchange (IKE) are the three main protocols used with IPsec.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers extending a network virtually.

*This page intentionally left blank*

CHAPTER 28

# Extending the Network Virtually

**This chapter covers the following official ENCOR 350-401 exam objectives:**

- 2.3 Describe network virtualization concepts
- 2.3.a LISP
- 2.3.b VXLAN

This chapter covers two topics that are critical for understanding the operation of SD-Access and SD-WAN: Locator ID/Separation Protocol (LISP) and Virtual Extensible LAN (VXLAN). LISP is a routing architecture that provides new semantics for IP addressing. Traditionally, the IP address of an endpoint would denote its location and its identity. Using the same value for both location and identity limits the security and management of networks. LISP routing architecture design separates the device identity, or endpoint identifier (EID), from its location or routing locator (RLOC). This provides the benefit of simplified multihoming and facilitates scalable any-to-any WAN connectivity.

The limitation of LISP is that it only supports Layer 3 overlays. It does not carry MAC addresses because it discards the Layer 2 Ethernet header. Because MAC addresses need to be carried in certain fabric technologies, such as SD-Access, you need VXLAN. VXLAN is a tunneling protocol that allows for the tunneling of Ethernet traffic over an IP network. You can use VXLAN to deploy a Layer 2 overlay network on top of a Layer 3 underlay network. This preserves the original Layer 2 header and addresses the limitation of LISP.

**This chapter covers the following technology topics:**

- Locator ID/Separation Protocol (LISP)
- Virtual Extensible LAN (VXLAN)

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. What does Cisco' SD-Access solution use as its control plane?
2. What is the name for the VXLAN unique network ID?

### Answers

1. LISP
2. VXLAN network identifier (VNI)

# Locator ID/Separation Protocol (LISP)

Cisco Locator/ID Separation Protocol (LISP) is a routing architecture that is a simple, incrementally deployable, network-based solution. It does the following:

- Enables enterprise and service providers to simplify multihomed routing
- Allows for highly scalable network virtualization
- Aids in reducing operational complexities

The challenge is that the current Internet routing and addressing architecture uses a single numbering space, the IP address, to simultaneously express two attributes about a device: its identity and its location. One very visible and detrimental result of this single numbering space is manifested in the rapid growth of the Internet's default-free zone (DFZ). The growth of the DFZ is due to multihoming, non-aggregatable address allocations, and business events such as mergers and acquisitions.

The solution that LISP offers is the ability to split the device identity, known as an endpoint identifier (EID), and its location, known as its routing locator (RLOC), into two different numbering spaces. Splitting the EID and RLOC functions allows for many benefits, including these:

- Simplified and cost-effective multihoming
- Ingress traffic engineering (TE) capabilities

- IP address and host mobility, including session persistence across mobility events
- IPv6 transition simplification, including incremental deployment of IPv6 using existing IPv4 infrastructure
- Simplified highly scalable network virtualization

Let's take a high-level look at the services that are enabled by using LISP:

- Improved routing system scalability through the use of topologically aggregated RLOCs
- Provider independence for devices that are numbered outside the EID space
- Multihoming of end sites with improved traffic engineering
- IPv6 transition functionality

LISP is deployed primarily in network edge devices. It requires no changes to host stacks, DNS, or local network infrastructure, and it requires few major changes to existing network infrastructures.

ExamAlert

Before taking the ENCOR exam, make sure you understand the components that make up a LISP architecture.

The following architecture components make up LISP:

- **Endpoint identifier (EID):** An EID is assigned to an end host and is the IP address of the endpoint within the LISP site. It is the same IP address in use on endpoints (IPv4 and IPv6); the IP address on an endpoint and an EID operate the same.
- **LISP site:** A LISP site is a site where LISP routers and EIDs reside.
- **Ingress tunnel router (ITR):** An ITR receives packets from endpoints within a LISP site and either encapsulates packets to remote LISP sites or natively forwards packets to non-LISP sites.
- **Egress tunnel router (ETR):** An ETR receives packets from interfaces facing the network core and either decapsulates LISP packets or natively delivers non-LISP packets to local EIDs at the site.

- **Tunnel router (xTR):** An xTR is a router that performs the functions of both an ITR and an ETR.
- **Proxy ITR (PITR):** A PITR is a LISP infrastructure device that provides connectivity between non-LISP sites and LISP sites. A PITR does this by advertising coarse-aggregate prefixes for the LISP EID namespace into the Internet, thus attracting non-LISP traffic destined to LISP sites. The PITR is responsible for encapsulating packets in a non-LISP site that are headed to a LISP site.
- **Proxy ETR (PETR):** A PETR is a LISP infrastructure device that allows IPv6 LISP sites without native IPv6 RLOC connectivity to reach LISP sites that only have IPv6 RLOC connectivity. The PETR is responsible for decapsulating the LISP packets that arrive from a LISP site and are headed to a non-LISP site.
- **Proxy xTR (PxTR):** It is likely that a LISP interworking device will implement both PITR and PETR functions. When this is the case, the device is referred to as a PxTR.
- **LISP router:** A LISP router is a router that performs the functions of any or all of the following: ITR, ETR, PITR, and/or PETR.
- **Routing locator (RLOC):** An RLOC is an IPv4 or IPv6 address assigned to a device (typically a router) in the global routing system. An RLOC is an IPv4 or IPv6 address of an ETR that decapsulates LISP packets.
- **Map server (MS):** An MS configures the LISP site policy for LISP ETRs that register to it. This includes the EID prefixes for which registering ETRs are authoritative and an authentication key, which must match the one that is configured on the ETR.
- **Map resolver (MR):** An MR receives map requests encapsulated to it from ITRs, and when configured with a service interface to the LISP alternative topology (ALT), it forwards map requests to the ALT. Map resolvers also send negative map replies to ITRs in response to queries for non-LISP addresses.
- **Map server/map resolver (MS/MR):** When the map server and map resolver functions are implemented on the same device, it is referred to as an MS/MR.

Figure 28.1 provides a very general overview of the LISP deployment environment. It depicts the essential areas that exist in a LISP environment: the LISP sites (EID namespaces) and the non-LISP sites (RLOC namespaces) and the LISP infrastructure.

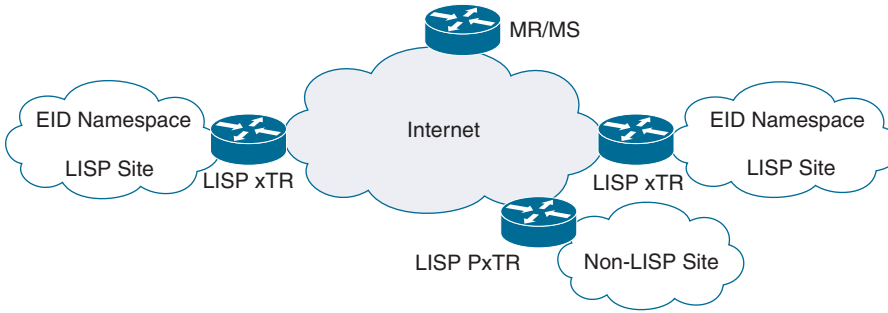



FIGURE 28.1 **LISP Deployment Environment**

# LISP Architecture

The LISP architecture has three main components:

- **LISP routing architecture:** LISP separates the IP addresses into EIDs and RLOCs. This enables endpoints to roam from site to site, and the only thing that changes is their RLOCs. The EIDs remain the same.
- **LISP control plane:** The goal of the control plane is to determine where to send traffic. It works similarly to Domain Name System (DNS). As DNS can resolve a name to an IP address, LISP can resolve an EID to an RLOC by sending map requests to the MR. LISP uses map requests and map replies, which are similar to DNS requests and replies. LISP uses a distributed mapping system to map EIDs to RLOCs. When an ITR needs to find an RLOC address, a map request is sent to the mapping system. This makes the system an efficient and scalable on-demand routing protocol because it is based on a pull model instead of the push model used with traditional routing protocols.
- **LISP data plane:** After an ITR has determined which RLOC to use to reach an EID, it needs to encapsulate the IP packets. The original IP header (known as the inner header) and data are preserved. The encapsulation involves the following three headers:
  - **LISP header:** This IP header includes LISP information on how to forward the IP packet. One field in the LISP header that is worth mentioning is the instance ID. The instance ID is a 24-bit unique identifier that keeps prefixes separately when there are overlapping EID addresses in a LISP site.

- **Outer LISP UDP header:** This header contains the source port that an ITR selects to prevent traffic from one LISP site to another taking the same path to the destination—even if they are equal-cost multipathing (ECMP) links. The destination port that UDP uses here is 4341.

- **Outer LISP IP header:** This header contains the source and destination RLOC IP addresses needed to route a packet from an ITR to an ETR.

Next, let's take a look at the path of a LISP packet:

- The source endpoint performs a DNS lookup to find the destination.
- Traffic is remote, so traffic is sent to the branch router.
- The branch router does not know how to get to the specific address of the destination. It is LISP enabled, and it performs a LISP lookup to find a locator address. The LISP mapping database informs the branch router how to get to the one or more available addresses to get the packet to the destination. The LISP mapping database can return priority and weight as part of this lookup to help with traffic engineering and shaping.
- The branch router performs an IP-in-IP encapsulation and transmits the data out of the appropriate interface based on standard IP routing decisions.

The receiving LISP-enabled router receives the packet, decapsulates the packet, and forwards the packet to the final destination.

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which device in a LISP deployment receives map requests that are encapsulated to it from ITRs?
   - ❍ **A.** Map server (MS)
   - ❍ **B.** Map resolver (MR)
   - ❍ **C.** Alternative topology (ALT)
   - ❍ **D.** Egress tunnel router (ETR)

2. Which of the following statements best describes endpoint identifiers (EIDs)?
   - ❍ **A.** EIDs are assigned to the physical interface of routers.
   - ❍ **B.** EIDs are assigned to routers.
   - ❍ **C.** EIDs are assigned to loopback interfaces.
   - ❍ **D.** EIDs are assigned to end hosts.

3. Which of the following is true of an endpoint identifier (EID)?
   - ❍ **A.** Can be IPv4 or IPv6
   - ❍ **B.** Can be IPv4 only
   - ❍ **C.** Can be IPv6 only
   - ❍ **D.** Can be the MAC address only

# Answers

1. **B** is correct. The map resolver (MR) is deployed as a LISP infrastructure device. It receives map requests that are encapsulated to it from ingress tunnel routers (ITRs).
2. **D** is correct. An EID is assigned to an end host and is the IP address of an endpoint within a LISP site.
3. **A** is correct. EIDs are the same IP addresses (IPv4 and IPv6) in use on endpoints.

# Virtual Extensible LAN (VXLAN)

**ExamAlert**

Before taking the ENCOR exam, make sure you understand why VXLAN was created and what problem it solves.

LISP is fantastic, but it is unable to preserve Layer 2 Ethernet headers. In environments where you need to preserve Layer 2 header information, you can use VXLAN. Like LISP, VXLAN is an overlay technology for network virtualization. However, it provides a Layer 2 extension over a shared Layer 3 underlay infrastructure network, using MAC-in-IP or UDP tunneling encapsulation. The goals of a Layer 2 extension in the overlay network are to overcome the limitations of geographic boundaries and achieve flexibility for workload placement within or across a data center or site. Each of these overlays is termed a *VXLAN segment*.

As the name *VXLAN* implies, the technology is meant to provide the same services to connected Ethernet end systems that VLANs provide today—but in a more extensible manner. VXLAN is extensible in terms of the scale and the reach of a deployment. Whereas the 802.1Q VLAN ID space is only 12 bits, the VXLAN ID space is 24 bits. This doubling in size allows the ID space to increase from 4094 VLANs to over 16 million unique VXLAN identifiers and should provide sufficient room for expansion for years to come.

VXLAN uses IP (both unicast and multicast) as the transport medium. The ubiquity of IP networks and equipment means you can extend the end-to-end reach of a VXLAN segment far beyond the typical reach of VLANs using 802.1Q today. Each VXLAN segment can be highly distributed among the networking nodes. With so many segments, the number of end systems connected to any one segment is expected to be relatively low. Therefore, the percentage of network nodes that participate in any one segment will also be low.

Figure 28.2 shows the 8-byte VXLAN header that consists of a 24-bit VXLAN network identifier (VNI) and a few reserved bits. This VXLAN header along with the original Ethernet frame goes in the UDP payload. The 24-bit VNI is used to identify Layer 2 segments and to maintain Layer 2 isolation between the segments.

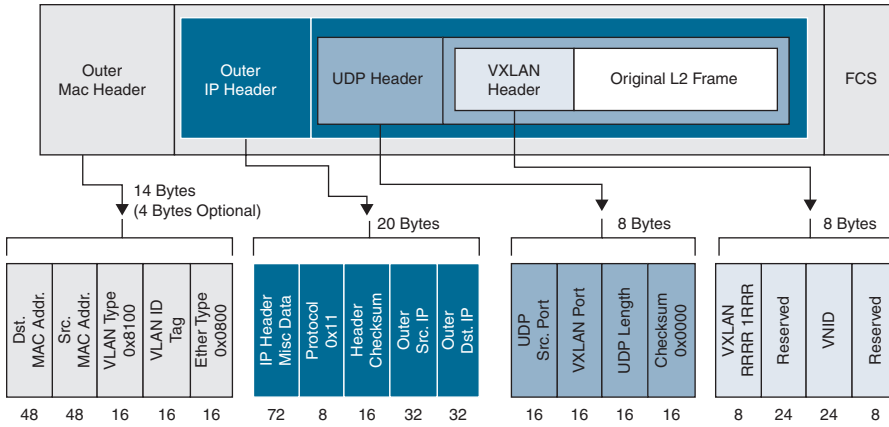



FIGURE 28.2 **VXLAN Packet Format**

Some of the high-level benefits of VXLAN are as follows:

- **Highly dynamic end systems:** End systems connected using VXLAN can be very dynamic in terms of creation, deletion, power-on and power-off, and mobility across the network nodes.
- **Integration with existing widely deployed network equipment:** VXLAN can be used with Ethernet switches and IP routers. A single administrative domain is used to administer the network infrastructure. This aspect is consistent with operations within a data center and not across the Internet.
- **Low network node overhead and simple implementation:** With the requirement to support very large numbers of network nodes, the resource requirements on each node should not be excessive in terms of memory footprint or processing cycles. This requirement also considers hardware offload.

Before looking into VXLAN overlays, let's recap the types of overlays:

- **Layer 2 overlays:** Layer 2 overlays emulate Layer 2 LAN segments. The forwarding process uses Ethernet frame headers and can therefore transport both IP and non-IP packets. The overlay creates a single subnet (that is, a single Layer 2 domain). Layer 2 overlays are helpful in an environment where workload mobility is necessary. A Layer 2 overlay can emulate a particular physical topology. The disadvantage of Layer 2 overlays is that the Layer 2 domain becomes wider; therefore, the flooding domain becomes bigger as well.

- **Layer 3 overlays:** Layer 3 overlays are helpful when you want to transport IP packets and you also want to abstract the underlying IP infrastructure. Layer 3 overlays are handy when you need IP mobility. You can also use them to carry IPv4 traffic over an IPv6 infrastructure or IPv6 traffic over an IPv4 infrastructure.

With VXLAN, Layer 2 overlay networks are carried/tunneled over a Layer 3 underlay network.

The three drivers for VXLAN in traditional networks are greater scalability, isolation for multitenancy, and any-to-any Layer 2 communications. VXLAN provides a unique network ID called a VXLAN (or virtual) network identifier (VNI), which provides a tunnel for transporting the original payload. The VNI tunnel (or overlay) keeps the traffic separate from other VXLAN traffic, much like what happens with virtual routing and forwarding (VRF). This approach provides multitenancy and therefore security and scale.

With VXLAN, you create LAN segments using an overlay approach with MAC-in-UDP encapsulation and a 24-bit segment identifier in the form of a VXLAN ID. Because VXLAN uses UDP or IP encapsulation, you can apply ECMP and load balance network workload, using all available paths with the same route metric.

Some of the key characteristics of VXLAN technology and overlays are as follows:

- VXLAN uses MAC-in-UDP/IP encapsulation.
- VXLAN can work on hypervisor-based virtual switches and physical switches.
- VXLAN uses ECMP to achieve optimal path usage over the transport network.

Of importance in a VXLAN infrastructure is the use of VXLAN tunnel endpoints (VTEPs). A VTEP is an entity that originates or terminates a VXLAN tunnel. The VTEP encapsulates the frame and sends it through the transport network to the destination VTEP. That VTEP decapsulates the frame and sends it to the final destination. From the transport network perspective, you can use any IP-capable device to send a VXLAN-encapsulated Layer 2 frame.

Each VTEP has two interfaces:

- One interface provides a bridging function for local hosts (which can be a trunk port to the access switch).

- The other interface has an IP identification in the core network for VXLAN encapsulation and decapsulation.

Each VTEP has one or more IP addresses. These addresses work as source IP addresses in the packet that contains the original Layer 2 frame and goes through the transport network. The encapsulation carries the VXLAN identifier to isolate each Layer 2 network from any other Layer 2 network. Platforms such as Cisco Catalyst 9300 support VTEPs.

Figure 28.3 shows the functional components of VTEPs and the logical topology that is created for Layer 2 connectivity across the transport IP network.

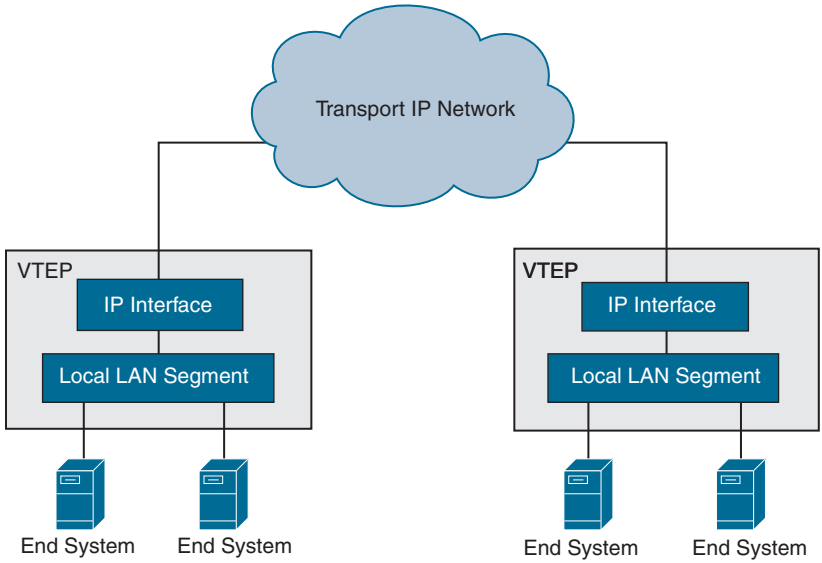



FIGURE 28.3 **VTEPs**

VXLAN is a data plane protocol that can be used with other control plane protocols. For example, in data centers and private cloud environments, MP-BGP and multicast are the most popular control planes used with VXLAN, whereas for the campus environment, the LISP control plane with VXLAN is preferred. These are the various control planes that are supported with VXLAN:

- VXLAN with LISP control plane
- VXLAN with MP-BGP EVPN control plane
- VXLAN with multicast underlay
- VXLAN with static unicast VXLAN tunnels

For campus deployment, Cisco's Software-Defined Access (SD-Access) is an example of an implementation of VXLAN with the LISP control plane. You can learn more about SD-Access in Chapter 23, "SD-Access."

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. In a VXLAN header, what is the size of the VXLAN ID field?
   - ❍ **A.** 12 bits
   - ❍ **B.** 20 bits
   - ❍ **C.** 24 bits
   - ❍ **D.** 48 bits

2. What kind of encapsulation does VXLAN use?
   - ❍ **A.** MAC-in-TCP
   - ❍ **B.** MAC-in-UDP
   - ❍ **C.** MAC-in-IPsec
   - ❍ **D.** MAC-in-ICMP

3. What is the name of the unique network identity provided by VXLAN?
   - ❍ **A.** VXLAN tunnel endpoint (VTEP)
   - ❍ **B.** Endpoint
   - ❍ **C.** VXLAN overlay
   - ❍ **D.** VXLAN network identifier (VNI)

## Answers

1. **C** is correct. The VXLAN ID space is 24 bits long.
2. **B** is correct. VXLAN creates LAN segments by using an overlay approach with the MAC-in-UDP encapsulation technique.
3. **D** is correct. VXLAN provides a unique network ID called a VXLAN (or virtual) network identifier (VNI), which provides a tunnel for transporting the original payload.

# Review Questions

1. Which LISP device provides connectivity when traffic is flowing from a non-LISP site to a LISP site?
   - ❍ **A.** Ingress tunnel router (ITR)
   - ❍ **B.** Egress tunnel router (ETR)
   - ❍ **C.** Proxy ITR (PITR)
   - ❍ **D.** Proxy ETR (PETR)

2. Which of the following describes a routing locator (RLOC)?
   - ❍ **A.** The IPv4 or IPv6 address of the customer router that is decapsulating a LISP packet
   - ❍ **B.** A logical topology that is deployed as part of the LISP infrastructure to provide scalable EID prefix aggregation
   - ❍ **C.** A block of IP addresses that are assigned to a site
   - ❍ **D.** The name of a site where LISP routers and EIDs reside

3. Which of the following is the preferred deployment scenario in the campus?
   - ❍ **A.** VXLAN with LISP control plane
   - ❍ **B.** VXLAN with MP-BGP EVPN control plane
   - ❍ **C.** VXLAN with multicast underlay
   - ❍ **D.** VXLAN with static unicast VXLAN tunnels

4. How many unique identifiers does VXLAN provide?
   - ❍ **A.** 12
   - ❍ **B.** 24
   - ❍ **C.** 400,000
   - ❍ **D.** 16,000,000

# Answers to Review Questions

1. **C** is correct. PITR is a LISP infrastructure device that provides connectivity between non-LISP sites and LISP sites.
2. **A** is correct. RLOC is an IPv4 or IPv6 address assigned to a device (typically a router) in the global routing system.
3. **A** is correct. For the campus environment, LISP control plane with VXLAN is preferred.
4. **D** is correct. The 24-bit VXLAN ID space allows for 16,000,000 unique identifiers.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*
- *CCNP and CCIE Data Center Core DCCOR 350-601 Official Cert Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers troubleshooting.

CHAPTER 29

# Troubleshooting

**This chapter covers the following official ENCOR 350-401 exam objective:**

- 4.1 Diagnose network problems using tools such as debugs, conditional debugs, trace route, ping, SNMP, and syslog

This chapter covers concepts related to network assurance. This chapter is divided into two sections. The first section looks at the various IOS tools that are used to troubleshoot and monitor a network. It looks at common use cases and operations of diagnostic tools such as **debug**, **traceroute**, and **ping**. The second section of this chapter examines how Simple Network Management Protocol (SNMP) exposes the environment and performance parameters of a network device. It looks at how SNMP configurations allow a network management system (NMS) to collect and process data. This chapter does not address the syslog portion of ENCOR 350-401 Objective 4.1, as that is covered in Chapter 30, "Monitoring."

**This chapter covers the following technology topics:**

- Troubleshooting Overview
- Simple Network Management Protocol (SNMP)

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. What debugging feature allows you to selectively enable debugging and logging for specific features?
2. What command is used to discover the routes that a router's packets follow on their way to a destination?
3. Which extended **ping** option is useful in determining the smallest MTU in the path to a destination?
4. What component of SNMP is the system that controls and monitors the activities of the network hosts?

### Answers

1. Conditional debugging
2. **traceroute**
3. Do Not Fragment (DF) bit
4. SNMP manager

# Troubleshooting Overview

Network engineers are often called upon to troubleshoot networks. The troubleshooting may involve investigating a problem and then following through with diagnosis and resolution. This section looks at the tools and techniques that are commonly part of a network engineer's toolkit and how these tools and techniques can be used to troubleshoot small to large-scale networks. This section covers the following common IOS tools and provides examples of their usage:

- **debug:** The **debug** command helps in isolating protocol and configuration problems.
- **traceroute:** The **traceroute** command provides a method to determine the route that packets take to reach their destination from one device to another.
- **ping:** The **ping** command is used to test Layer 3 end-to-end connectivity and can aid in the isolation of protocol and configuration problems.

# Using debug to Analyze Traffic

The **debug** command can provide real-time information about the traffic being seen (or not seen) on an interface, error messages generated by hosts on the network, protocol-specific diagnostic packets, and other useful troubleshooting data.

Output formats vary with each **debug** command. Some generate a single line of output per packet, and others generate multiple lines of output per packet. Some generate large amounts of output, and others generate only occasional output. Some generate lines of text, and others generate information in field format.

Care should be taken to use **debug** commands to isolate problems and not to monitor normal network operations. The high processor overhead of running **debug** commands can disrupt the operation of a router or switch. You should use **debug** commands only when looking for specific types of traffic or problems and when you have already narrowed the issues to a likely subset of causes.

Many **debug** commands are processor intensive and can cause serious network problems, including degraded performance or loss of connectivity. This is especially true if they are enabled on an already heavily loaded router. When you finish using a **debug** command, it is critical to remember to disable it with its specific **no debug** command or use the **no debug all** command or **undebug all** command to turn off all debugging.

Running **debug** in a production environment can be especially problematic if the router CPU is not powerful enough and there are a lot of process-switched packets. These circumstances can easily lead to the router stalling, and **debug** should be used with caution. One way of minimizing the impact of the **debug** command on a router is to use an access control list (ACL) to narrow down the specific traffic that should be monitored. For example, one way of using the **debug** command with an ACL is to use the command **debug ip packet** [**access-list** *number*]. This way, only packets that match the access list criteria will be subject to **debug ip packet**. In this case, the access list is applied to the debug operation and not to an interface.

Example 29.1 shows the use of an ACL with the **debug** command. In this example, a continuous **ping** is being sent from the remote host (100.1.1.1) to a router (100.1.1.2). Note the 100 matches on the configured ACL. This ensures that the **debug ip packet** command only debugs packets from host 100.1.1.2 to host 100.1.1.1. Packets that do not match this ACL are not debugged.

EXAMPLE 29.1 **Using the debug Command and an ACL**

```
R2# debug ip packet ?
  <1-199>      Access list
  <1300-2699>  Access list (expanded range)
  detail       Print more debugging detail
  <cr>         <cr>

R2#
R2# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R2(config)# access-list 100 permit ip host 100.1.1.2 host 100.1.1.1
R2(config)# end
R2#
R2# debug ip packet 100
IP packet debugging is on for access list 100
R2# show debug
Generic IP:
  IP packet debugging is on for access list 100

R2# show access-list
Extended IP access list 100
    10 permit ip host 100.1.1.2 host 100.1.1.1 (100 matches)
```

Another way of minimizing the **debug** command's impact is to buffer the **debug** messages and show them by using the **show logging** command after **debug** has been turned off. However, there is a downside to logging to a buffer: If the buffer is not set to a high enough value, the debugging messages can get overwritten with new debugging messages. However, if the buffer is set too high, you may be taking away resources from other services and features that need memory on the device. It is a balancing act.

Example 29.2 shows how to use the **debug** command to buffer debug messages and the **show logging** command to view the messages.

EXAMPLE 29.2 **Buffering Debug Messages**

```
R2#
R2# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R2(config)# no logging console
R2(config)# logging buffer 5000
R2(config)# end

R2#
R2# debug ip packet
```

```
IP packet debugging is on
R2# ping 100.1.1.1 source 100.1.1.2
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 100.1.1.1, timeout is 2 seconds:
Packet sent with a source address of 100.1.1.2
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 6/8/11 ms
R2# undebug all
All possible debugging has been turned off
R2# show logging
Syslog logging: enabled (0 messages dropped, 3 messages rate-limited,
0 flushes, 0 overruns, xml disabled, filtering disabled)

No Active Message Discriminator.

No Inactive Message Discriminator.


    Console logging: disabled
    Monitor logging: level debugging, 0 messages logged, xml disabled,
                     filtering disabled
    Buffer logging:  level debugging, 284 messages logged, xml
disabled,
                    filtering disabled
    Exception Logging: size (8192 bytes)
    Count and timestamp logging messages: disabled
    Persistent logging: disabled

No active filter modules.

    Trap logging: level informational, 126 message lines logged
        Logging Source-Interface:       VRF Name:

Log Buffer (5000 bytes):
E, mtu 0, fwdchk FALSE

*Mar  4 18:22:13.430: IP: s=100.1.1.2 (local), d=100.1.1.1 (Giga-
bitEthernet0/0), len 100, sending
*Mar  4 18:22:13.430: IP: s=100.1.1.2 (local), d=100.1.1.1 (Giga-
bitEthernet0/0), len 100, sending full packet
*Mar  4 18:22:13.436: IP: s=100.1.1.1 (GigabitEthernet0/0),
d=100.1.1.2, len 100, input feature, MCI Check(109), rtype 0, forus
FALSE, sendself FALSE, mtu 0, fwdchk FALSE
*Mar  4 18:22:13.437: IP: tableid=0, s=100.1.1.1 (GigabitEthernet0/0),
d=100.1.1.2 (Loopback0), routed via RIB
*Mar  4 18:22:13.437: IP: s=100.1.1.1 (GigabitEthernet0/0),
d=100.1.1.2, len 100, rcvd 4
*Mar  4 18:22:13.437: IP: s=100.1.1.1 (GigabitEthernet0/0),
d=100.1.1.2, len 100, stop process pak for forus packet
```

**ExamAlert**

Before taking the ENCOR exam, make sure you understand the use of the conditional **debug** command to selectively enable debugging.

A third way of narrowing down the output of a **debug** command is by using the conditional **debug** command. When a conditional **debug** command is enabled, the output is only generated for packets that contain the information specified in the configured condition.

The conditional debugging feature allows you to selectively enable debugging and logging for specific features, based on the set of conditions you define. Conditional **debug** commands are especially useful for allowing granular debugging in a network operating on a large scale with many features. They allow you to observe detailed debugging for granular instances within the system. This is useful when you need to debug only a particular session among thousands of sessions. It is also possible to specify multiple conditions when doing conditional debugging.

A condition that is being debugged can be a feature or an identity, which could be an interface, an IP address, or a MAC address. In contrast, the general **debug** command produces output without discriminating the feature objects that are being processed. A general **debug** command consumes a lot of system resources and impacts system performance. Conditional **debug** commands help you mitigate the risk of a router crashing when running the general **debug** command.

Example 29.3 shows conditional debugging of OSPF adjacency on a particular interface.

EXAMPLE 29.3 **Conditional Debugging**

```
R2# debug condition ?
  called       called number
  calling      calling
  cpl          Cisco Provisioning Language debugging
  glbp         interface group
  interface    interface
  ip           IP address
  mac-address  MAC address
  match-list   apply the match-list
  profile      Media Services Profile
  standby      interface group
  username     username
  vcid         VC ID
  vrf          Virtual Routing and Forwarding
  xconnect     Xconnect conditional debugging on segment pair
```

```
R2# debug ip ospf adj
OSPF adjacency debugging is on
R2# debug condition interface g0/0
Condition 1 set
R2#
*Mar  4 21:23:41.931: OSPF-1 ADJ   Gi0/0: Cannot see ourself in hello
from 10.10.10.1, state INIT
*Mar  4 21:23:46.690: OSPF-1 ADJ   Gi0/0: Rcv DBD from 10.10.10.1 seq
0xF49 opt 0x52 flag 0x7 len 32  mtu 1500 state INIT
*Mar  4 21:23:46.690: OSPF-1 ADJ   Gi0/0: 2 Way Communication to
10.10.10.1, state 2WAY
*Mar  4 21:23:46.690: OSPF-1 ADJ   Gi0/0: Nbr 10.10.10.1: Prepare
dbase exchange
*Mar  4 21:23:46.690: OSPF-1 ADJ   Gi0/0: Send DBD to 10.10.10.1 seq
0x257F opt 0x52 flag 0x7 len 32
*Mar  4 21:23:46.690: OSPF-1 ADJ   Gi0/0: First DBD and we are not
SLAVE
*Mar  4 21:23:46.699: OSPF-1 ADJ   Gi0/0: Rcv DBD from 10.10.10.1 seq
0x257F opt 0x52 flag 0x2 len 112  mtu 1500 state EXSTART
*Mar  4 21:23:46.699: OSPF-1 ADJ   Gi0/0: NBR Negotiation Done. We are
the MASTER
*Mar  4 21:23:46.699: OSPF-1 ADJ   Gi0/0: Nbr 10.10.10.1: Summary list
built, size 4
*Mar  4 21:23:46.699: OSPF-1 ADJ   Gi0/0: Send DBD to 10.10.10.1 seq
0x2580 opt 0x52 flag 0x1 len 72
*Mar  4 21:23:46.706: OSPF-1 ADJ   Gi0/0: Rcv LS REQ from 10.10.10.1
length 36 LSA count 1
*Mar  4 21:23:46.707: OSPF-1 ADJ   Gi0/0: Send LS UPD to 10.10.10.1
length 100 LSA count 1
*Mar  4 21:23:46.709: OSPF-1 ADJ   Gi0/0: Rcv DBD from 10.10.10.1 seq
0x2580 opt 0x52 flag 0x0 len 32  mtu 1500 state EXCHANGE
*Mar  4 21:23:46.709: OSPF-1 ADJ   Gi0/0: Exchange Done with
10.10.10.1
```

In this example, although the interface of multiple neighboring routers was shut down and brought back online, the conditional **debug** captured only the OSPF adjacency change that occurred on interface g0/0.

# Troubleshooting with traceroute

The **traceroute** command discovers the routes that a router's packets follow on their way to a destination. On the other hand, an extended **traceroute** command allows you to specify IP header options, which means a router can perform a more extensive range of tests.

The **traceroute** command works by using the error messages generated by routers when a datagram exceeds its Time-to-Live (TTL) value. The device that executes the **traceroute** command sends out a sequence of User Datagram

Protocol (UDP) datagrams, with incrementing TTL values, to an invalid port address (which is, by default, 33434) at the remote host. First, probe datagrams are sent with a TTL value of 1. This causes the first router to discard the probe datagrams and send back ICMP time exceeded messages. The **traceroute** command then sends several probes and displays the round-trip time for each. After every third probe, the TTL is increased by 1.

The **traceroute** command terminates when the destination responds that the maximum TTL is exceeded or when the user interrupts the trace with the escape sequence. It is a good idea to use the **traceroute** command when the network is functioning correctly to see how the command works under normal conditions; this gives you a baseline for comparison during troubleshooting.

Table 29.1 shows the possible characters in **traceroute** output.

TABLE 29.1 **traceroute Output Characters**

| Character | Description |
|---|---|
| **nn msec** | For each host, the round-trip time, in milliseconds, for the specific number of probes |
| ***** | The probe timed out |
| **A** | Administratively prohibited (for example, an access list) |
| **Q** | Source quench (that is, destination too busy) |
| **I** | User interrupted test |
| **U** | Port unreachable |
| **H** | Host unreachable |
| **N** | Network unreachable |
| **P** | Protocol unreachable |
| **T** | Timeout |
| **?** | Unknown packet type |

Example 29.4 shows the use of the **traceroute** command. IP packets are sent to the destination address (10.10.30.2 in this case) with a TTL value that increments up to the maximum (30 by default). Each router in the destination path decrements the TTL field by 1 unit as it forwards the packet. When a router in the path finds that the TTL is 1, it responds to the source with an ICMP "time exceeded" message. This lets the source know that the packet traversed that particular router as a hop.

EXAMPLE 29.4 **Using the traceroute Command**

```
R1#
R1# traceroute 10.10.30.2
Type escape sequence to abort.
Tracing the route to 10.10.30.2
VRF info: (vrf in name/id, vrf out name/id)
  1 10.10.10.2 3 msec 3 msec 3 msec
  2 10.10.20.2 4 msec 4 msec 4 msec
  3 10.10.30.2 5 msec *  7 msec
R1#
```

Example 29.5 shows the output of an extended **traceroute** command. This variation of the **traceroute** command allows you to see what paths packets take to reach the destination, help you troubleshoot routing loops, or help you determine when packets are getting lost. This is useful when packets are being blocked by an ACL.

EXAMPLE 29.5 **Using an Extended traceroute Command**

```
R1#
R1# traceroute
Protocol [ip]:
Target IP address: 10.10.30.2
Ingress traceroute [n]:
Source address: 10.10.10.1
Numeric display [n]:
Timeout in seconds [3]:
Probe count [3]:
Minimum Time to Live [1]:
Maximum Time to Live [30]: 5
Port Number [33434]:
Loose, Strict, Record, Timestamp, Verbose[none]:
Type escape sequence to abort.
Tracing the route to 10.10.30.2
VRF info: (vrf in name/id, vrf out name/id)
  1 10.10.10.2 3 msec 3 msec 3 msec
  2 10.10.20.2 5 msec 4 msec 6 msec
  3 10.10.30.2 10 msec *  5 msec
R1#
```

You can use both **ping** and **traceroute** commands to do a minimal performance test on a link. You can use the **ping** and **traceroute** commands to obtain the round-trip time (that is, the time required to send an echo packet and get an answer back). This can be useful as it gives you a rough idea of the delay on the link. However, these figures should be used with caution because they are

not precise enough for performance evaluation. When a packet destination is a router, this packet has to be process switched. The processor has to handle the information from this packet and send a response. Because this is not a router's primary goal, answering a **ping** is offered as a best-effort service.

An extended **traceroute** command is a variation of the **traceroute** command. An extended **traceroute** command can be used to see what path packets take to get to a particular destination. The command can be used to check routing at the same time. This is helpful when you are troubleshoot routing loops or when you are determining where packets are getting lost (for example, if a route is missing, if packets are being blocked by a firewall, or if packets are filtered by an ACL).

It is common to use an extended **ping** command to determine the type of connectivity problem and then follow up by using the extended **traceroute** command to narrow down the location of the problem occurrence.

These are some of the common messages that the **traceroute** command can return:

- **time exceeded:** This error message indicates that an intermediate communication server has seen and discarded the packet.
- **destination unreachable:** This error message indicates that the destination node has received the probe and discarded it because it could not deliver the packet. If the timer goes off before a response comes in, **traceroute** prints an asterisk (*****).

The **traceroute** command terminates when any of the following occurs:

- The destination responds.
- The maximum TTL is exceeded.
- The user interrupts the trace with the escape sequence (that is, **Ctrl+Shift+6**).

Table 29.2 highlights the fields that need to be entered with an extended **traceroute** command.

TABLE 29.2 **Extended traceroute Fields**

| Field | Description |
|---|---|
| **Protocol [ip]:** | Prompts for a supported protocol. The default is IP. |
| **Target IP address:** | Prompts for the IP address or hostname of the destination host you plan to ping. |

| Field | Description |
|---|---|
| **Source address:** | The interface or IP address of the router to use as the source address of the probes. |
| **Numeric display [n]:** | The default is to have both symbolic and numeric displays. However, the symbolic display can be suppressed. |
| **Timeout in seconds [3]:** | The number of seconds to wait for a response to a probe packet. The default is 3 seconds. |
| **Probe count [3]:** | The number of probes to be sent at each TTL level. The default count is 3. |
| **Minimum Time to Live [1]:** | The TTL value for the first probes. The default value is 1 but can be set to a higher value to suppress the display of known hops. |
| **Maximum Time to Live [30]:** | The largest TTL value that can be used. The default is 30. The **traceroute** command terminates when the destination is reached or this value is reached. |
| **Port Number [33434]:** | The destination port used by UDP probe messages. The default is 33434. |
| **Loose, Strict, Record, Timestamp, Verbose [none]:** | IP header options. Although the **traceroute** command places the requested options in each probe, there is no guarantee that all routers will process the options. |

# Troubleshooting with ping

The **ping** command is used to check host reachability and network connectivity. It can be invoked from both user EXEC mode and privileged EXEC mode. The **ping** command can be used to verify connectivity for AppleTalk, IP, and other protocols.

For IP, the **ping** command sends Internet Control Message Protocol (ICMP) echo messages. ICMP is the protocol that reports errors and provides information relevant to IP packet addressing. When a host receives an ICMP echo message, it sends an ICMP echo reply message back to the source. As with **traceroute**, it is good to look at the output from the **ping** command when the network is operating correctly and not just when it is broken. This helps you see the command output under normal conditions so that you can have something to compare with when troubleshooting.

ICMP echo messages can be used to determine the following:

- Whether a remote host is active or inactive
- The round-trip delay in communicating with the host
- Packet loss

The **ping** command sends an echo request packet to an address and then waits for a reply. The **ping** is successful only if both of the following occur:

- The echo request gets to the destination.
- The destination is able to get an echo reply back to the source within a predetermined time, called a timeout. (The default timeout value on a Cisco router is 2 seconds.)

Table 29.3 shows the possible characters in the **ping** command's output.

TABLE 29.3 **ping Output Character**

| Character | Description |
|---|---|
| **!** | Receipt of a reply |
| **.** | Network server timed out while waiting for a reply |
| **U** | A destination unreachable error PDU was received |
| **Q** | Source quench (that is, destination too busy) |
| **M** | Could not fragment |
| **?** | Unknown packet type |
| **&** | Packet lifetime exceeded |

When a normal **ping** command is sent from a router to a destination, the source address of the **ping** is the IP address of the interface that the packet used to exit the router. When an extended **ping** command is used, the source IP address can be changed to any IP address on the router. An extended **ping** command allows you to do more advanced checks of host reachability and network connectivity. It works only in privileged EXEC mode, whereas a normal **ping** command works at both user and privileged modes. To use an extended **ping** command, you type **ping** at the command line, press **Enter**, and fill in the fields as you're prompted to do so.

Example 29.6 shows an example of the **ping** command. In this case, the **ping** command sends an echo request packet to the address 10.10.30.2 and waits for a reply. As you can see in the output, five request packets were sent, and five were successfully received, as indicated by the exclamation point. (A period would indicate that there was a timeout while waiting for a reply.) The output **Success rate is 100 percent** shows the percentage of packets that successfully echoed back to the router. Anything less than 80% is usually considered problematic. The output **Round-trip min/avg/max = 8/10/12 ms** shows the round-trip travel time intervals for the protocol echo packets, including minimum, average, and maximum (in milliseconds).

EXAMPLE 29.6 **Using the ping Command**

```
R1# ping 10.10.30.2
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 10.10.30.2, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 8/10/12 ms
R1#
```

Example 29.7 shows an example of the **ping** command with a repeat count of 100. By default, the repeat count is 5. If you need more than 5 echo request packets, you can change the repeat count. This might be useful, for example, if you need to send additional echo request packets to a destination while you work on troubleshooting an issue. As the issue is resolved, you may start to see successful ICMP echo replies.

EXAMPLE 29.7 **Using the ping Command to Repeat Count**

```
R1# ping 10.10.30.2 repeat 100
Type escape sequence to abort.
Sending 100, 100-byte ICMP Echos to 10.10.30.2, timeout is 2 seconds:
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!
Success rate is 100 percent (100/100), round-trip min/avg/max =
4/12/22 ms
R1#
```

Example 29.8 shows an example of the **ping** command with a size of 1500. To determine whether the default size of 1500 byes is working, you can test with the DF bit set between a source and a destination. In this case, you are testing MTU size of 1500 bytes to destination 10.10.30.2.

EXAMPLE 29.8 **Using the ping Command with the Size Specified**

```
R1#
R1# ping 10.10.30.2 size 1500
Type escape sequence to abort.
Sending 5, 1500-byte ICMP Echos to 10.10.30.2, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 7/9/12 ms
R1#
```

Table 29.4 highlights the fields that need to be entered with an extended **ping**.

TABLE 29.4 **Extended ping Fields**

| Field | Description |
|---|---|
| **Protocol [ip]:** | Prompts for a supported protocol. The default is IP. |
| **Target IP address:** | Prompts for the IP address or hostname of the destination host you plan to **ping**. |
| **Repeat count [5]:** | Number of **ping** packets that are sent to the destination. The default is 5. |
| **Datagram size [100]:** | Size of the **ping** packet, in bytes. The default is 100 bytes. |
| **Timeout in seconds [2]:** | Timeout interval. The default is 2 seconds. The **ping** is declared successful only if the echo reply packet is received before this time interval. |
| **Extended commands [n]:** | Specifies whether a series of additional commands appears. The default is no. |
| **Source address or interface:** | The interface or IP address of the router to use as the source address of the probes. |
| **Type of service [0]:** | Specifies the Type of Service (ToS). The requested ToS is placed in each probe, but there is no guarantee that all routers process the ToS. The default is 0. |
| **Set DF bit in IP header? [no]:** | Specifies whether the don't fragment (DF) bit is set on the **ping** packet. This is useful in determining the smallest MTU in the path to a destination. The default is no. |
| **Validate reply data? [no]:** | Specifies whether to validate the reply data. The default is no. |
| **Data pattern [0xABCD]** | Specifies the data pattern. Different data patterns are used to troubleshoot framing errors and clocking problems on serial lines. The default is 0xABCD. |
| **Loose, Strict, Record, Timestamp, Verbose [none]:** | The default IP header option is none. These are the options:<br>• **Loose** allows you to influence the path by specifying the address(es) of the hop(s) you want the packet to go through.<br>• **Strict** is used to specify the hop(s) that you want the packet to go through with no other hop(s) allowed to be visited.<br>• **Record** displays the address(es) of the hops (up to nine) that the packet goes through.<br>• **Timestamp** is used to measure round-trip time to a particular host.<br>• **Verbose** is automatically selected along with any other option. |
| **Sweep range of sizes [n]:** | Allows you to vary the sizes of the echo packets that are being sent. This is used to determine the minimum sizes of the MTUs configured on the nodes along the path to the destination address. Performance problems caused by packet fragmentation are thus reduced. The default is no. |

**ExamAlert**

For the ENCOR exam, make sure you understand why you may want to use an extended **ping** while troubleshooting.

It is common to use extended **ping** commands to quickly test connectivity by sending different sizes of packets to a destination. For example, you might need to send 1500-byte packets with the DF bit set to make sure there are no MTU issues on the interfaces or to test different QoS policies that restrict certain packet sizes.

Example 29.9 shows an example of an extended **ping** command. When the normal **ping** command is sent from a router, the source address of the **ping** is the IP address of the interface that the packet uses to exit the router. With extended **ping**, the source IP address can be changed to any IP address on the router. The extended **ping** is used to perform a more advanced check of host reachability and network connectivity. As you can see in this example, the extended version of this command gives you more granular control of the fields that can be set.

EXAMPLE 29.9 **Using an Extended ping Command**

```
R1# ping
Protocol [ip]:
Target IP address: 10.10.30.2
Repeat count [5]: 1
Datagram size [100]: 1500
Timeout in seconds [2]: 1
Extended commands [n]: yes
Ingress ping [n]:
Source address or interface: 10.10.10.1
DSCP Value [0]:
Type of service [0]:
Set DF bit in IP header? [no]: yes
Validate reply data? [no]:
Data pattern [0x0000ABCD]:
Loose, Strict, Record, Timestamp, Verbose[none]: Timestamp
Number of timestamps [ 9 ]: 3
Loose, Strict, Record, Timestamp, Verbose[TV]:
Sweep range of sizes [n]:
Type escape sequence to abort.
Sending 1, 1500-byte ICMP Echos to 10.10.30.2, timeout is 1 seconds:
Packet sent with a source address of 10.10.10.1
```

```
Packet sent with the DF bit set
Packet has IP options:  Total option bytes= 16, padded length=16
 Timestamp: Type 0.  Overflows: 0 length 16, ptr 5
  >>Current pointer<<
  Time= 00:00:00.000 UTC (00000000)
  Time= 00:00:00.000 UTC (00000000)
  Time= 00:00:00.000 UTC (00000000)
Reply to request 0 (16 ms).  Received packet has options
 Total option bytes= 16, padded length=16
 Timestamp: Type 0.  Overflows: 5 length 16, ptr 17
  Time=*17:47:24.698 UTC (83D13E9A)
  Time=*17:42:36.307 UTC (83CCD813)
  Time=*17:42:52.509 UTC (83CD175D)
  >>Current pointer<<

Success rate is 100 percent (1/1), round-trip min/avg/max = 16/16/16
ms
R1#
```

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which command can be used to turn off all debugging on a router?
   - ❍ **A. no ip debug**
   - ❍ **B. debug all**
   - ❍ **C. undebug all**
   - ❍ **D. undebug**

2. What character in the output of the **traceroute** command indicates that the packet has been dropped due to being filtered by an access list?
   - ❍ **A. A**
   - ❍ **B. I**
   - ❍ **C. U**
   - ❍ **D. N**

3. For what reason would you not want to use **ping** and **traceroute** output to gather precise performance information for a link when the destination is the router itself?
   - ❍ **A.** The packet is fast switched.
   - ❍ **B.** The packet is process switched.
   - ❍ **C.** The packet is CEF switched.
   - ❍ **D.** The **ping** command cannot be used to get any link performance information.

4. What is the maximum number of hops that can be specified with the **traceroute** command?
   - ❍ **A.** 10
   - ❍ **B.** 15
   - ❍ **C.** 20
   - ❍ **D.** 30

# Answers

1. **C** is correct. The **undebug** command all or **no debug all** command can be used to turn off all debugging.
2. **A** is correct. The **A** character in the output of the **traceroute** command means that the packet is administratively prohibited (for example, by an access list).
3. **B** is correct. Both the **ping** and **traceroute** commands do a minimal performance test on the link by obtaining the round-trip time. These figures should be used with care because they are not precise enough to be used for performance evaluation, especially when the destination is the router itself (because the packet has to be process switched).
4. **D** is correct. The largest TTL value that can be specified with the **traceroute** command is 30. This is also the default value.

# Simple Network Management Protocol (SNMP)

Simple Network Management Protocol (SNMP) is an application-layer protocol that provides a message format for communication between SNMP managers and agents. SNMP provides a standardized framework and a common language for monitoring and managing devices in a network. It is a simple and easy-to-implement protocol and is supported by a lot of vendors. It uses UDP as its transport mechanism to retrieve and send management information such as management information base (MIB) variables. It is common to use SNMP to gather information and performance data such as CPU usage, memory usage, interface traffic, and so on.

**ExamAlert**

Before taking the ENCOR exam, make sure you understand the different components of SNMP and what they are used for.

The SNMP framework has the following components:

- **SNMP manager:** The SNMP manager is a system that controls and monitors the activities of the network hosts that are using SNMP. The most common managing system is a network management system (NMS). The main job of the SNMP manager or NMS is to collect management data from managed devices via polling or trap messages. The SNMP manager periodically polls the SNMP agents on the managed devices by querying the devices for data. The main disadvantage of periodically polling for data is that there is a delay between an event's occurrence and the time when the SNMP manager polls the data.
- **SNMP agent:** The SNMP agent is the software component within a managed Cisco IOS network device that maintains the data for the device and reports this data, as needed, to the NMS. The agent resides on devices including routers, switches, and servers. To enable an SNMP agent on these devices, a relationship needs to be defined between the manager and the agent. Agents also generate SNMP traps, which are unsolicited notifications sent from an agent to a manager. SNMP traps are event based and provide almost real-time event notifications.
- **SNMP MIB:** The MIB is a repository (local database) for information about device parameters. An SNMP agent contains MIB variables, whose values the SNMP manager can request or change through Get or Set operations. An SNMP manager can get a value from an agent or store

a value in that agent, which sits in the Cisco IOS network device. The agent gathers data from the SNMP MIB. The agent can also respond to the manager requests to get or set data.

SNMP performs the following operations to retrieve data, modify SNMP object variables, and send the notification:

- **SNMP Get:** The NMS performs the Get operation to retrieve SNMP object variables. There are three Get operations:
  - **Get:** This operation retrieves the exact object from the SNMP agent.
  - **Get Next:** This operation retrieves the next object variables.
  - **Get Bulk:** This operation retrieves a large amount of object variable data without the need for repeated Get Next operations.
- **SNMP Set:** The NMS performs the SNMP Set operation to modify the value of an object variable.
- **SNMP traps:** A key feature of SNMP is its capability to generate unsolicited notifications, called traps, from SNMP agents.

Figure 29.1 illustrates the process of retrieving data through SNMP.

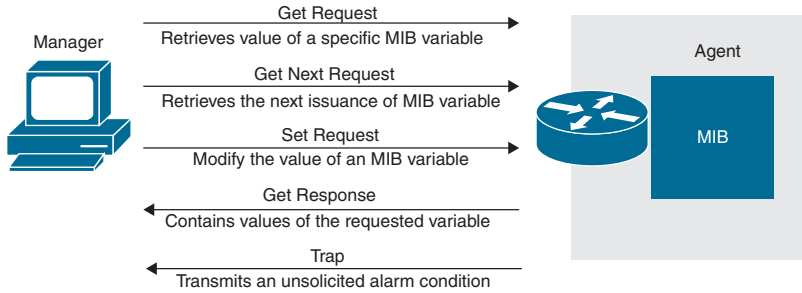



FIGURE 29.1 **Retrieving Data Through SNMP**

*Notification* is considered a generic term covering both traps and informs:

- **Traps:** These messages alert the SNMP manager to a condition on the network. Traps are less reliable than informs because the receiver does not send an acknowledgment when it receives a trap.
- **Informs:** These messages are acknowledged traps. An inform is a trap that includes a request for confirmation of receipt from the SNMP manager. The SNMP agent has no way of knowing if the SNMP manager received an SNMP trap. An SNMP inform request packet is sent continually until an SNMP acknowledgment is received.

There are three versions of SNMP:

- **SNMPv1:** Security is based on a community string. SNMPv1 introduced five message types: Get Request, Get Next Request, Set Request, Get Response, and Trap.
- **SNMPv2c:** Security is based on a community string. Two new message types were introduced with SNMPv2: Get Bulk Request and Inform Request. SNMPv2c includes support for a bulk retrieval mechanism and detailed error message reporting to management stations. This bulk retrieval mechanism supports the retrieval of tables and large quantities of information, minimizing the number of round trips required. SNMPv2c also improved error handling support by including expanded error codes that distinguish the different types of errors.
- **SNMPv3:** SNMPv3 provides secure access to devices by providing authentication, encryption, integrity, authorization, and access control. With SNMPv3, an authentication strategy is set up for a user and the group in which the user resides. A security level is the permitted level of security within a security model. With SNMP authentication and access control, a security model identifies the SNMP version that is associated with a user for the group in which the user resides. A combination of a security model and a security level determines which security mechanism is employed when handling an SNMP packet.

SNMPv3 introduced three levels of security:

- **noAuthNoPriv:** Authentication occurs using a string match of the username. This is in plaintext and is configured between the NMS and the network device. No encryption is provided.
- **authNoPriv:** Authentication is based on a Message Digest 5 (MD5) or Secure Hash Algorithm (SHA) hash, and no encryption is provided; messages are sent in plaintext.
- **authPriv:** In addition to authentication using MD5 or SHA hash, CBC-DES encryption is used to encrypt the messages between the NMS and the network device.

Table 29.5 shows the combinations of security models and levels and their meanings.

TABLE 29.5 **SNMP Security Model and Levels**

| | Level | Authentication | Encryption | Description |
|---|---|---|---|---|
| **SNMPv1** | | | | |
| | noAuthNoPriv | Community String | No | Uses a community string match for authentication |
| **SNMPv2c** | | | | |
| | noAuthNoPriv | Community String | No | Uses a community string match for authentication |
| **SNMPv3** | | | | |
| | noAuthNoPriv | Username | No | Uses a username match for authentication |
| | authNoPriv | Message Digest 5 (MD5) or Secure Hash Algorithm (SHA) | No | Provides authentication based on the HMAC-MD5 or HMAC-SHA algorithms |
| | authPriv | MD5 or SHA | Data Encryption Standard (DES), 3DES, and AES | Provides authentication based on the HMAC-MD5 or HMAC-SHA algorithms. In addition to authentication, it provides encryption based on the CBC-DES (DES-56), 3DES, and AES-256 standard. |

Example 29.10 shows a basic SNMPv2c configuration and verification. This example uses the public read-only community string and enables support for EIGRP notification. The switch is then enabled to send inform requests to host 10.10.10.99 using the string defined as public with SNMPv2c. The command **show snmp host** shows the configured variables.

EXAMPLE 29.10 **Configuring SNMPv2c**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# snmp-server community public RO
SW1(config)# snmp-server enable traps eigrp
SW1(config)# snmp-server host 10.10.10.99 version 2c public
SW1(config)# end
SW1# show snmp host
Notification host: 10.10.10.99  udp-port: 162   type: trap
user: public    security model: v2c
SW1#
```

Example 29.11 shows a basic SNMPv3 configuration. Similarly to the previous example, this example permits any SNMP manager to access all objects with read-only permissions using the named public community string. This example does not cause the device to send traps. user1 is configured to receive traps at the noAuthNoPriv security level when the SNMPv3 security model is enabled. Then user2 is configured to receive traps at the authNoPriv security level when the SNMPv3 security model is enabled. Finally, user3 is configured to receive traps at the priv security level when the SNMPv3 security model is enabled.

EXAMPLE 29.11 **Configuring SNMPv3**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# snmp-server community public RO
SW1(config)# snmp-server group group1 v3 noauth
SW1(config)# snmp-server user user1 group1 remote 10.10.10.99
SW1(config)# snmp-server host 10.10.10.99 informs version 3 noauth
user1 config
SW1(config)# snmp-server group group2 v3 auth
SW1(config)# snmp-server user user2 group2 remote 10.10.10.99 v3 auth
md5 ExamCramPassword1
SW1(config)# snmp-server group group3 v3 priv
SW1(config)# snmp-server user user3 group3 remote 10.10.10.99 v3 auth
md5 ExamCramPassword1 priv access des56
SW1(config)# end
SW1#
```

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. What is an SNMP trap message used for?

   - ❍ **A.** To transmit an unsolicited message about a condition on a device from the agent to an NMS
   - ❍ **B.** Nothing; a trap is not an SNMP message
   - ❍ **C.** To perform a write to the MIB variable on the agent
   - ❍ **D.** For requesting confirmation of receipt from the SNMP manager

2. What is an SNMP inform message used for?
   - ❍ **A.** To transmit an unsolicited message about a condition on a device from the agent to an NMS
   - ❍ **B.** To reliably alert the SNMP manager to a condition on the network
   - ❍ **C.** To perform a write to the MIB variable on the agent
   - ❍ **D.** To request confirmation of receipt from the SNMP manager

3. Which SNMP-related components are supported on a Cisco IOS device? (Choose two.)
   - ❍ **A.** NMS
   - ❍ **B.** SNMP manager
   - ❍ **C.** SNMP agent
   - ❍ **D.** MIB

# Answers

1. **A** is correct. Traps are unsolicited messages that alert the SNMP manager to a conditions on network devices.
2. **B** is correct. With SNMP informs, an SNMP inform request packet is sent continually until an acknowledgment is received. This makes informs more reliable than traps.
3. **C** and **D** are correct. The SNMP agent is the software component within a managed Cisco IOS network device that maintains the data for the device and reports this data, as needed, to the NMS. An SNMP agent contains MIB variables, whose values the SNMP manager can request or change through Get or Set operations. An SNMP manager can get a value from an agent or store a value in that agent that sits in the Cisco IOS network device.

# Review Questions

1. Which of the following are ways to minimize the impact of running the **debug ip packet** command on a router? (Choose two.)
   - ❍ **A.** Use an ACL to narrow the traffic that should be monitored
   - ❍ **B.** Run the **debug** command in user EXEC mode only
   - ❍ **C.** Buffer the **debug** messages and view them from log afterward
   - ❍ **D.** Run the **debug** command in privileged EXEC mode only

2. What character in the output of a **traceroute** command indicates that the network is unreachable?
   - ❍ **A. A**
   - ❍ **B. I**
   - ❍ **C. U**
   - ❍ **D. N**

3. Which of the following cannot be determined by using the **ping** command?
   - ❍ **A.** Whether a remote host is active or inactive
   - ❍ **B.** The round-trip delay in communicating with the host
   - ❍ **C.** The route that a packet follows on the way to a destination
   - ❍ **D.** Packet loss

4. Which of the following is not a valid SNMPv3 security level?
   - ❍ **A.** noAuthNoPriv
   - ❍ **B.** privNoAuth
   - ❍ **C.** authPriv
   - ❍ **D.** authNoPriv

# Answers to Review Questions

1. **A** and **C** are correct. One way of minimizing the impact of the **debug** command on a router is to use an access list to narrow down the specific traffic that should be monitored. A second way of minimizing the impact of the **debug** command is to buffer the **debug** messages and show them by using the **show log** command after **debug** has been turned off.
2. **D** is correct. The character **N** from a **traceroute** output indicates that the network is unreachable.

3. **C** is correct. The **ping** command cannot be used to determine a path that packets follow from source to destination node. This is done using the **traceroute** command.
4. **B** is correct. noAuthNoPriv, privNoAuth, authPriv, and authNoPriv are all security levels that are part of SNMPv3.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers monitoring.

*This page intentionally left blank*

CHAPTER 30

# Monitoring

**This chapter covers the following official ENCOR 350-401 exam objectives:**

- 4.1 Diagnose network problems using tools such as debugs, conditional debugs, trace route, ping, SNMP, and syslog
- 4.2 Configure and verify device monitoring using syslog for remote logging
- 4.3 Configure and verify NetFlow and Flexible NetFlow
- 4.4 Configure and verify SPAN/RSPAN/ERSPAN

This chapter focuses on the syslog section of ENCOR 350-401 Objective 4.1. (Chapter 29, "Troubleshooting," covers the other sections of this objective.) This chapter takes a look at managing the messages that Cisco devices generate. It looks at configuring settings related to messages sent to the console, the logging buffer, and a remote syslog collector. It also looks at Cisco NetFlow and Flexible NetFlow and how you can use them to gain visibility in a network. Finally, this chapter looks at the different implementations of switch port analyzer services, such as Switch Port Analyzer (SPAN), Remote SPAN (RSPAN), and Encapsulated Remote SPAN (ERSPAN).

**This chapter covers the following technology topics:**

- Syslog
- NetFlow and Flexible NetFlow
- Switch Port Analyzer (SPAN), Remote SPAN (RSPAN), and Encapsulated Remote SPAN (ERSPAN)

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. Which logging severity level should be used with caution because it can hamper the network's performance?
2. What are the different components of Flexible NetFlow?
3. Which SPAN supports source ports, source VLANs, and destinations across different switches?

### Answers

1. Debugging (level 7)
2. Flow record, flow monitor, flow exporter, and flow sampler
3. Encapsulated Remote SPAN (ERSPAN)

# Syslog

Most Cisco network devices use the syslog protocol to manage system logs and alerts. Some Cisco devices lack the storage space and RAM needed to store logs locally (as servers do). To overcome these limitations, Cisco devices have two options:

- **Internal buffer:** With an internal buffer, the system allows a small portion of memory buffer to log the most recent messages. This logging option is enabled by default. However, using the internal buffer for logging has a major disadvantage: The logged messages are lost when the device reboots.
- **Syslog:** The UNIX-style syslog protocol can send messages to a remote device for storage. Redirection of logs to an external syslog server is not enabled by default. The significant advantage of using a remote syslog server is that there is no dependency on the storage space of the local device. The only limitation in terms of log storage space is the available disk space on the external syslog server. The logs stored on a remote server serve as a backup and are useful for merging and log analysis.

The focus of this section is on syslog server configuration. However, logs can also be configured to be displayed on the console and terminal line as well as being part of Simple Network Management Protocol (SNMP) traps. With console logging, a device sends all logs to the console port. Users physically connected to the console port can view these messages. With terminal logging, which is similar to console logging, log messages are displayed to the device vty lines. Finally, a device can use SNMP traps to send log messages to an external SNMP server.

System log messages can have up to 80 characters and a percentage sign (%), as well as an optional sequence number or timestamp information, if configured. Messages are displayed in the following format:

*seq no:timestamp: %facility-severity-MNEMONIC:description*

Table 30.1 describes the syslog message elements.

TABLE 30.1 **Syslog Message Elements**

| Element | Description |
|---|---|
| *seq no:* | Stamps log messages with a sequence number only if the **service sequence-numbers** command is configured. |
| *timestamp* | Indicates the date and time of the message event. It appears only if the **service timestamps log** command is configured. Logs are displayed in UTC time by default. |
| *facility* | Indicates the facility to which a message refers (for example, SNMP, SYS). A facility can be a hardware device, a protocol, or a module of the system software. It indicates the sources and cause of the system message. |
| *severity* | Indicates the severity of the message, on a scale from 0 to 7. |
| *MNEMONIC* | Uniquely describes the message. |
| *description* | Contains detailed information about the event being reported. |

Before you enable logging, it is critical to ensure that a device that will be sending logging information to a syslog server is configured with accurate date and time information. If it is not configured appropriately, the timestamps of all the logging messages will not reflect the appropriate and accurate times—and that would make troubleshooting harder because you would not be able to correlate issues with the logs by using the timestamps generated. A straightforward way of addressing this issue is to have Network Time Protocol (NTP) configured properly so that devices receive the correct time.

Messages generated by a device have specific severity levels associated with them, and these severity levels and what is logged can be changed as needed. The severity level indicates how important the message is. For example, you may be more interested in seeing a message related to an interface flap than in seeing a message about a network administrator successfully logging on to switch to carry out a routine task.

ExamAlert

For the ENCOR exam, make sure you are familiar with the various syslog severity levels.

Table 30.2 shows the default log severity level of each message type.

TABLE 30.2 **Syslog Severity Levels**

| Log Level Keyword | Level | Description | Syslog Definition |
|---|---|---|---|
| **emergencies** | 0 | System unstable | **LOG_EMERG** |
| **alerts** | 1 | Immediate action needed | **LOG_ALERT** |
| **critical** | 2 | Critical conditions | **LOG_CRIT** |
| **errors** | 3 | Error conditions | **LOG_ERR** |
| **warnings** | 4 | Warning conditions | **LOG_WARNING** |
| **notifications** | 5 | Normal but significant condition | **LOG_NOTICE** |
| **informational** | 6 | Informational messages only | **LOG_INFO** |
| **debugging** | 7 | Debugging messages | **LOG_DEBUG** |

You can limit the logs sent to a syslog server based on severity level by using the **logging trap** *level* command in global configuration mode. To limit logging message to the console by type, you use the **logging console** *level* command. To limit logging message to the terminal lines by type, you use the **logging monitor** *level* command. Finally, to limit logging message to the local buffer by type, you use the **logging buffer** *level* command. If severity level 0 is configured, only emergency messages will be displayed. However, if severity level 3 is configured, all messages up to level 3 (emergency, alert, critical, and error) will be displayed.

One crucial point that should be considered when selecting the logging level is the use of logging severity level 7 (debugging messages). Logging severity level 7 should be used with caution as it can present too much information and hamper network performance.

It is possible to set a limit on the number of messages that a device logs per second. You can enable the limit for all messages sent to the console, and you can specify that messages of a specific severity are exempt from the limit. To enable a logging rate limit, you use the **logging rate-limit** command in global configuration mode.

Example 30.1 shows the configuration of logging to buffer severity levels.

EXAMPLE 30.1 **Configuring Logging to Buffer Severity Level**

```
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# logging buffer ?
  <0-7>              Logging severity level
  <4096-2147483647>  Logging buffer size
  alerts             Immediate action needed           (severity=1)
  critical           Critical conditions               (severity=2)
  debugging          Debugging messages                (severity=7)
  discriminator      Establish MD-Buffer association
  emergencies        System is unusable                (severity=0)
  errors             Error conditions                  (severity=3)
  filtered           Enable filtered logging
  informational      Informational messages            (severity=6)
  notifications      Normal but significant conditions (severity=5)
  warnings           Warning conditions                (severity=4)
  xml                Enable logging in XML to XML logging buffer
  <cr>               <cr>
R1(config)# exit
R1#
```

Example 30.2 shows the configuration and verification of logging to a syslog server. The **logging host 10.10.10.99** command specifies that you are logging the syslog server at 10.10.10.99. Specifying the trap level 3 in the command **logging trap 3** means that you are logging with severity levels 0, 1, 2, and 3. The command **logging rate-limit 10** specifies logging at 10 logs per second.

EXAMPLE 30.2 **Configuring Syslog**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# logging host 10.10.10.99
R1(config)# logging trap 3
R1(config)# logging rate-limit 10
R1(config)# exit
R1# show logging
```

```
Syslog logging: enabled (0 messages dropped, 3 messages rate-limited,
0 flushes, 0 overruns, xml disabled, filtering disabled)

No Active Message Discriminator.

No Inactive Message Discriminator.


    Console logging: disabled
    Monitor logging: level debugging, 0 messages logged, xml disabled,
                     filtering disabled
    Buffer logging:  level debugging, 80 messages logged, xml
disabled,
                    filtering disabled
    Exception Logging: size (8192 bytes)
    Count and timestamp logging messages: disabled
    Persistent logging: disabled

No active filter modules.

    Trap logging: level errors, 83 message lines logged
        Logging to 10.10.10.99  (udp port 514, audit disabled,
              link up),
              4 message lines logged,
              0 message lines rate-limited,
              0 message lines dropped-by-MD,
              xml disabled, sequence number disabled
              filtering disabled
        Logging Source-Interface:       VRF Name:
```

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which logging severity level indicates that the system is unstable?
   - ❍ **A.** Emergency
   - ❍ **B.** Warning
   - ❍ **C.** Error
   - ❍ **D.** Critical

2. Which logging severity level indicates a normal but significant condition?
   - ❍ **A.** Emergency
   - ❍ **B.** Notification
   - ❍ **C.** Informational
   - ❍ **D.** Critical

3. All except which of the following are elements of a syslog message?
   - ❍ **A.** Sequence number
   - ❍ **B.** Timestamp
   - ❍ **C.** Severity
   - ❍ **D.** Notification

# Answers

1. **A** is correct. Logging severity level 0 (emergency) indicates that the system is unstable.
2. **B** is correct. Logging severity level 5 (notification) indicates a normal but significant condition.
3. **D** is correct. Notification is not an element of a syslog message, but it is a severity level.

# NetFlow and Flexible NetFlow

Visibility into the network is an indispensable ability for network engineers. Due to ongoing new requirements and pressures, it is critical that network engineers understand how a network is behaving. NetFlow is embedded in Cisco IOS software to help characterize network operation. It helps network engineers understand the following network behaviors:

- Application and network usage
- Network productivity and utilization of network resources
- The impact of changes to the network
- Network anomaly and security vulnerabilities
- Long-term compliance issues

Cisco NetFlow enables network engineers to understand the who, what, when, where, and how of network traffic flows. When you use NetFlow to understand network behavior, business processes improve, and you have access to an audit trail showing how the network is used. This increased awareness reduces the vulnerability of the network to outages and allows for efficient operation of the network.

The ability to characterize IP traffic and understand how and where it flows is critical for network availability, performance, and troubleshooting. Monitoring IP traffic flows facilitates more accurate capacity planning and ensures that resources are used appropriately, where needed. It helps a network engineer determine where to apply quality of service (QoS) and how to optimize resource usage. It plays a vital role in network security to detect denial-of-service (DoS) attacks, network-propagated worms, and other undesirable network events.

NetFlow provides several mechanisms to address many common problems encountered on networks:

- **Analysis of new applications and their network impact:** NetFlow enables you to analyze new application network loads due to voice over IP (VoIP) or the addition of remote sites.
- **Reduction of peak WAN traffic:** NetFlow statistics can measure WAN traffic improvement from application policy changes to help you understand who is using the network and determine the top talkers on the network.

- **Troubleshooting and understanding of network pain points:** NetFlow can help you diagnose slow network performance, bandwidth hogs, and bandwidth utilization quickly with the command-line interface or reporting tools.
- **Detection of unauthorized WAN traffic:** NetFlow can be used to avoid costly upgrades by identifying applications that are causing congestion.
- **Security and anomaly detection:** NetFlow can be used for anomaly detection and worm diagnosis.
- **Validation of QoS parameters:** NetFlow can be used to confirm that the appropriate bandwidth has been allocated to each class of service (CoS) and that no CoS is oversubscribed or undersubscribed.

NetFlow provides information by examining each packet forwarded within a router or switch for a set of IP packet attributes. These attributes are the packet's IP packet identity or fingerprint and determine whether the packet is unique or similar to other packets.

The IP flow can be based on a set of five and up to seven of the following IP packet attributes:

- IP source address
- IP destination address
- Source port
- Destination port
- Layer 3 protocol type
- Class of service
- Router or switch interface

All packets with the same source/destination IP address, source/destination ports, protocol interface, and CoS are grouped into a flow, and then packets and bytes are tallied. This methodology of fingerprinting or determining a flow is scalable because a large amount of network information is condensed into a database of NetFlow information called the *NetFlow cache*.

The flow information is useful for understanding these network behaviors:

- The source address helps you understand who is originating the traffic.
- The destination address tells you who is receiving the traffic.

- Ports characterize the application using the traffic.
- The class of service indicates the priority of the traffic.
- The device interface indicates how traffic is being utilized by the network device.
- Tallied packets and bytes show the amount of traffic.

Two methods can be used to access the data from NetFlow: the use of **show** commands at the command-line interface (CLI) and the use of an application reporting tool. If you are interested in an immediate view of what is happening in the network, the CLI can be used. The NetFlow CLI can also be beneficial for troubleshooting. The other choice is to export NetFlow to a reporting server (that is, a NetFlow collector). A NetFlow collector assembles and interprets the exported flows and combines them to produce valuable reports for traffic and security analysis. NetFlow periodically pushes information to the NetFlow reporting collector. The NetFlow cache is constantly filling with flows, and software in the router or switch searches the cache for flows that have terminated or expired. These flows are exported to the NetFlow collector server. Flows are terminated when the network communication has ended (that is, when a packet contains the TCP FIN flag).

You follow these steps to implement NetFlow data reporting:

1. Configure NetFlow to capture flows to the NetFlow cache.
2. Configure the NetFlow export to send flows to the collector.
3. Search the NetFlow cache for flows that have terminated and export them to the NetFlow collector server.
4. Bundle together approximately 30 to 50 flows and transport them, in User Datagram Protocol (UDP) format, to the NetFlow collector server.
5. Create real-time or historical reports from the data with the NetFlow collector software.

A flow is ready for export when it is inactive for a certain time (that is, when no new packets are received for the flow) or when the flow is long-lived (active) and lasts greater than the active timer (long FTP download, for example). Also, the flow is ready for export when a TCP flag indicates that the flow is terminated (that is, FIN or RST flag). There are timers to determine whether a flow is inactive or whether a flow is long-lived; the default for the inactive flow timer is 15 seconds and for the active flow timer is 30 minutes. All the timers for export are configurable.

**ExamAlert**

For the ENCOR exam, make sure you know the high-level steps for configuring both NetFlow and Flexible NetFlow.

Cisco NetFlow is configured on a per-interface basis. When it is configured on an interface, IP packet flow information is captured into the NetFlow cache. You can configure NetFlow on an interface by using the **ip flow** command. To export the NetFlow cache to a NetFlow collector, you need to first choose the version of the NetFlow export packet and then the collector's IP address.

Example 30.3 shows the configuration of NetFlow Version 9 and NetFlow data export on a router. In this case, the interface GigabitEthernet 0/0 is configured for NetFlow data capture and is exporting the data to the collector at 10.10.10.99. It is capturing both ingress and egress traffic.

EXAMPLE 30.3 **Configuring NetFlow**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# ip flow-export version 9
R1(config)# ip flow-export destination 10.10.10.99 9999
R1(config)# interface GigabitEthernet 0/0
R1(config-if)# ip flow ingress
R1(config-if)# ip flow egress
R1(config-if)# end
R1#
```

Example 30.4 shows the verification of NetFlow and NetFlow data export on a router. The **show ip flow interface** command shows the interfaces that are configured for NetFlow. The **show ip flow export** command shows the destination for the NetFlow data as well as statistics on the export, including any errors. The **show ip cache flow** command shows the traffic flow that NetFlow is capturing.

EXAMPLE 30.4 **Verifying NetFlow Configuration**

```
R1#
R1# show ip flow interface
GigabitEthernet0/0
  ip flow ingress
  ip flow egress
R1#
```

```
R1# show ip flow export
Flow export v9 is enabled for main cache
  Export source and destination details :
  VRF ID : Default
    Destination(1)  10.10.10.99 (9999)
  Version 9 flow records
  0 flows exported in 0 udp datagrams
  0 flows failed due to lack of export packet
  0 export packets were sent up to process level
  0 export packets were dropped due to no fib
  0 export packets were dropped due to adjacency issues
  0 export packets were dropped due to fragmentation failures
  0 export packets were dropped due to encapsulation fixup failures
R1#
R1# show ip cache flow
IP packet size distribution (48 total packets):
   1-32   64   96  128  160  192  224  256  288  320  352  384  416
448  480
   .000 .000 1.00 .000 .000 .000 .000 .000 .000 .000 .000 .000 .000
.000 .000

    512  544  576 1024 1536 2048 2560 3072 3584 4096 4608
   .000 .000 .000 .000 .000 .000 .000 .000 .000 .000 .000

IP Flow Switching Cache, 278544 bytes
  1 active, 4095 inactive, 1 added
  438 ager polls, 0 flow alloc failures
  Active flows timeout in 30 minutes
  Inactive flows timeout in 15 seconds
IP Sub Flow Cache, 34056 bytes
  1 active, 1023 inactive, 1 added, 1 added to flow
  0 alloc failures, 0 force free
  1 chunk, 1 chunk added
  last clearing of statistics never
Protocol    Total    Flows   Packets Bytes  Packets Active(Sec)
Idle(Sec)
--------    Flows    /Sec    /Flow  /Pkt   /Sec   /Flow     /Flow
SrcIf       SrcIPaddress   DstIf    DstIPaddress   Pr SrcP DstP  Pkts
Gi0/0       10.10.10.2     Null       224.0.0.5    59 0000 0000   48
R1#
```

**ip flow-top-talker** is a useful command for getting a quick snapshot of what is going on in a device from a flow perspective. It can be used with the **top** option for the number of talkers (1 through 200) and the **sort-by** option to sort by either bytes or packets.

Example 30.5 shows the configuration and verification of top talkers on a router. It shows the **ip flow-top-talkers** command used with **top 20** to show the top 20 talkers and the **sort-by bytes** command used to sort by bytes. Verification is done using the **show ip flow top-talkers** command.

EXAMPLE 30.5 **Configuring and Verifying Top Talkers**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# ip flow-top-talkers
R1(config-flow-top-talkers)# top 20
R1(config-flow-top-talkers)# sort-by bytes
R1(config-flow-top-talkers)# end
R1#
R1# show ip flow top-talkers

SrcIf         SrcIPaddress    DstIf         DstIPaddress    Pr SrcP
DstP Bytes
Gi0/0         10.10.10.2      Null          224.0.0.5       59 0000
0000    10K
1 of 20 top talkers shown. 1 flows processed.

R1#
```

Cisco IOS Flexible NetFlow is an extension to Cisco NetFlow Version 9 technology that allows for optimization of the network infrastructure, reduction of operation costs, improvement of capacity planning, and security incident detection with increased flexibility and scalability. Flexible NetFlow has many benefits beyond the traditional NetFlow functionality that has been available for years in Cisco hardware and software.

Flexible NetFlow allows for extremely granular and accurate traffic measurement and high-level aggregated traffic collection. Because it is part of Cisco IOS Software, Flexible NetFlow enables Cisco product-based networks to perform traffic flow analysis without the use of external probes, thereby making traffic analysis economical on large IP networks.

Some of the key advantages of Flexible NetFlow include the following:

- Flexibility and scalability of flow data beyond traditional NetFlow
- The ability to monitor a broader range of packet information, producing new information about network behavior
- Enhanced network anomaly and security detection

- User-configurable flow information to perform customized traffic identification and the ability to focus and monitor specific network behavior
- Convergence of multiple accounting technologies into one accounting mechanism

Traditional NetFlow tracks IP information such as IP addresses, ports, protocols, and TCP flags, whereas Flexible NetFlow can track a wide range of packet information for Layer 2, IPv4, and IPv6 flows, including the following:

- Source and destination MAC addresses
- Source and destination IPv4 or IPv6 addresses
- Source and destination TCP/UDP ports
- Type of service (ToS)
- Differentiated Services Code Point (DSCP)
- Packet and byte counts
- Flow timestamps
- Input and output interface numbers
- TCP flags and encapsulated protocol (TCP/UDP) and individual TCP flags
- Sections of packet for deep packet inspection
- All fields in the IPv4 header, including IP-ID and TTL
- All fields in the IPv6 header, including Flow Label and Option Header
- Routing information, such as next-hop address, source autonomous system (AS) number, destination AS number, source prefix mask, destination prefix mask, BGP next-hop, and BGP policy accounting traffic index

Flexible NetFlow consists of components that can be used together in several variations to perform traffic analysis and data export. The user-defined flow records and the component structure of Flexible NetFlow facilitate the creation of various configurations for traffic analysis and data export on a networking device with a minimum number of configuration commands. Each flow monitor can have a unique combination of a flow record, flow exporter, and cache type. If a change is made to a parameter such as the destination IP address for a flow exporter, it is automatically changed for all the flow monitors that use the flow exporter. The same flow monitor can be used in combination with different flow samplers to sample the same type of network traffic at different rates on different interfaces.

Flexible NetFlow has the following components:

- **Flow record:** Flow records are assigned to Flexible NetFlow flow monitors to define the cache used for storing flow data. Flexible NetFlow includes several predefined records. You can also define custom flow records.
- **Flow monitor:** Flow monitors are applied to interfaces to perform network traffic monitoring. Flow data is collected from the network traffic and added to the flow monitor cache during the monitoring process, based on the key and non-key fields in the flow record.
- **Flow exporter:** Flow exporters export the data in the flow monitor cache to a remote system, such as a server running NetFlow collector, for analysis and storage. Flow exporters are created as separate entities in the configuration. Flow exporters are assigned to flow monitors to provide data export capability for the flow monitors.
- **Flow sampler:** Flow samplers are created as separate components in a router's configuration. Flow samplers are used to reduce the load on the device that is running by limiting the number of packets selected for analysis.

Flexible NetFlow enables you to configure custom flow records for specific use cases; this is an extremely powerful capability.

Example 30.6 show the configuration and verification of a custom flow record. You will see how to create a custom flow exporter in a later example. This example shows how to create a custom flow record named **CUSTOM**. It uses the **match** command to match the IPv4 destination address. It uses the **collect** command to gather the byte and packet counts. For verification, it uses the **show flow record CUSTOM** command.

EXAMPLE 30.6 **Configuring and Verifying a Custom Flow Record**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# flow record CUSTOM
R1(config-flow-record)# description Custom Flow Record for IPv4
R1(config-flow-record)# match ipv4 destination address
R1(config-flow-record)# collect counter bytes
R1(config-flow-record)# collect counter packets
R1(config-flow-record)# end
R1#
R1# show flow record CUSTOM
```

```
flow record CUSTOM:
  Description:        Custom Flow Record for IPv4
  No. of users:       0
  Total field space:  12 bytes
  Fields:
    match ipv4 destination address
    collect counter bytes
    collect counter packets
```

Example 30.7 shows the configuration and verification of a custom flow exporter. The exporter is created to point to 10.10.10.99. It exports flow data from the device to a NetFlow collector.

EXAMPLE 30.7 **Configuring and Verifying Custom Flow Exporter**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# flow exporter CUSTOM
R1(config-flow-exporter)# description Export to Collector
R1(config-flow-exporter)# destination 10.10.10.99
R1(config-flow-exporter)# export-protocol netflow-v9
R1(config-flow-exporter)# transport UDP 9999
R1(config-flow-exporter)# end
R1#
R1# show flow exporter CUSTOM
Flow Exporter CUSTOM:
  Description:               Export to Collector
  Export protocol:           NetFlow Version 9
  Transport Configuration:
    Destination IP address: 10.10.10.99
    Source IP address:      10.10.10.1
    Transport Protocol:     UDP
    Destination Port:       9999
    Source Port:            49290
    DSCP:                   0x0
    TTL:                    255
    Output Features:        Not Used

R1#
```

Example 30.8 shows the configuration and verification of the flow monitor. In this example, the configured cache timeout value tells the router to export the cache to the NetFlow collection every 60 seconds.

EXAMPLE 30.8 **Configuring and Verifying Custom Flow Monitor**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# flow monitor CUSTOM
R1(config-flow-monitor)# description Uses Custom Flow Record CUSTOM
R1(config-flow-monitor)# record CUSTOM
R1(config-flow-monitor)# cache timeout active 60
R1(config-flow-monitor)# end
R1# show flow monitor CUSTOM
Flow Monitor CUSTOM:
  Description:       Uses Custom Flow Record CUSTOM
  Flow Record:       CUSTOM
  Flow Exporter:     CUSTOM (inactive)
  Cache:
    Type:                 normal
    Status:               not allocated
    Size:                 4096 entries / 0 bytes
    Inactive Timeout:     15 secs
    Active Timeout:       60 secs

R1#
```

One important piece in the configuration is the mapping of the flow exporter to the flow monitor. The flow exporter and the flow monitor need to be mapped together so that traffic that is being collected by the flow record can be exported to the NetFlow collector.

Example 30.9 shows the configuration and verification of the flow exporter mapping to the flow monitor. The flow exporter and the flow monitor need to be mapped so that traffic that is collected by the flow record can be exported to the NetFlow collector at 10.10.10.99.

EXAMPLE 30.9 **Configuring and Verifying the Flow Exporter Mapping to the Flow Monitor**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# flow monitor CUSTOM
R1(config-flow-monitor)# exporter CUSTOM
R1(config-flow-monitor)# end
R1#
R1# show flow monitor CUSTOM
Flow Monitor CUSTOM:
```

```
  Description:       Uses Custom Flow Record CUSTOM
  Flow Record:       CUSTOM
  Flow Exporter:     CUSTOM (inactive)
  Cache:
    Type:                 normal
    Status:               not allocated
    Size:                 4096 entries / 0 bytes
    Inactive Timeout:     15 secs
    Active Timeout:       60 secs

R1#
```

The final step in enabling Flexible NetFlow is to apply the flow monitor to an interface and turn on the collection of NetFlow statistics. This can be enabled for ingress or egress traffic or both.

Example 30.10 shows the configuration and verification of the flow monitor on an interface. This step enables Flexible NetFlow by applying the flow monitor on the interface GigabitEthernet 0/0. This example is configured for ingress operation using the **ip flow monitor CUSTOM** command. Verification is done using the **show flow monitor CUSTOM cache** command.

EXAMPLE 30.10 **Configuring and Verifying the Flow Monitor on an Interface**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# interface GigabitEthernet 0/0
R1(config-if)# ip flow monitor CUSTOM input
R1(config-if)# end
R1#
R1# show flow monitor CUSTOM cache
  Cache type:                               Normal
  Cache size:                                 4096
  Current entries:                               1
  High Watermark:                                1

  Flows added:                                   1
  Flows aged:                                    0
    - Active timeout      (    60 secs)          0
    - Inactive timeout    (    15 secs)          0
    - Event aged                                 0
    - Watermark aged                             0
    - Emergency aged                             0
```

```
IPV4 DST ADDR          bytes         pkts
===============  ===========  ===========
224.0.0.5                240            3

R1#
```

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which of the following is not an IP packet attribute that IP flow in NetFlow is based on?
   - ❍ **A.** IP source address
   - ❍ **B.** IP destination address
   - ❍ **C.** Source port
   - ❍ **D.** Destination MAC address

2. Which of the following can track a wide range of packet information for Layer 2, IPv4, and IPv6 flows?
   - ❍ **A.** NetFlow Version 9
   - ❍ **B.** NetFlow Version 10
   - ❍ **C.** Flexible NetFlow
   - ❍ **D.** Flexible NetFlow Version 7

3. All except which of the following are components of Flexible NetFlow?
   - ❍ **A.** Flow record
   - ❍ **B.** Flow session
   - ❍ **C.** Flow monitor
   - ❍ **D.** Flow exporter

## Answers

1. **D** is correct. The destination MAC address is not one of the packet attributes that IP flow is based on.
2. **C** is correct. Flexible NetFlow can track a wide range of packet information for Layer 2, IPv4, and IPv6 flows.
3. **B** is correct. Flow session is not a component of Flexible NetFlow.

# Switch Port Analyzer (SPAN), Remote SPAN (RSPAN), and Encapsulated Remote SPAN (ERSPAN)

SPAN copies traffic from one or more ports, one or more EtherChannels, or one or more VLANs and sends the copied traffic to one or more destinations for analysis by a network analyzer or network sniffer.

SPAN uses two different types of ports: a source port and a destination port. The source port is the port that is being monitored for traffic analysis. SPAN can be configured to copy ingress, egress, or both traffic types from the source port. Traffic is then copied to the destination port. The SPAN destination port (also called a monitoring port) receives a copy from the source ports and VLANs.

The destination port where a packet sniffer is plugged in has the following characteristics:

- A destination port must reside on the same switch as the source port (for a local SPAN session).
- A destination port can be any Ethernet physical port.
- A destination port can participate in only one SPAN session at a time. A destination port in one SPAN session cannot be a destination port for a second SPAN session.
- A destination port cannot be a source port.
- A destination port receives copies of sent and received traffic for all monitored source ports. If a destination port is oversubscribed, it can become congested. This congestion can affect traffic forwarding on one or more of the source ports.

A local SPAN session is an association of source ports and source VLANs with one or more destinations. Local SPAN copies traffic from one or more source ports in any VLAN or from one or more VLANs to a destination for analysis. A local SPAN session can only be configured on a local switch. Local SPAN does not have separate source and destination sessions.

When configuring a local SPAN session, if the traffic direction is not configured, the source sends both transmitted (Tx) and received (Rx) traffic to the destination port to be monitored. You can specify Tx, Rx, or both in the configuration.

Example 30.11 shows the configuration and verification of SPAN on a switch. The source interfaces for monitoring for session 1 are GigabitEthernet 0/0-1. Since the direction is not specified, it will monitor both Tx and Rx. The destination interface where the analyzer is plugged in is GigabitEthernet 0/2. The **show monitor session 1** command shows that this is a Local SPAN session, and it is monitoring in both directions; it also shows the source ports and the destination port for the session.

EXAMPLE 30.11 **Configuring and Verifying SPAN**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# monitor session 1 source interface
GigabitEthernet 0/0 - 1 ?
  ,     Specify another range of interfaces
  both  Monitor received and transmitted traffic
  rx    Monitor received traffic only
  tx    Monitor transmitted traffic only
  <cr>

SW1(config)# monitor session 1 source interface GigabitEthernet 0/0
- 1
SW1(config)# monitor session 1 destination  interface
GigabitEthernet 0/2
SW1(config)# exit
SW1#
SW1# show monitor session 1
Session 1
---------
Type                   : Local Session
Source Ports           :
    Both               : Gi0/0-1
Destination Ports      : Gi0/2
    Encapsulation      : Native
SW1#
```

**ExamAlert**

Before taking the ENCOR exam, make sure you understand the use case for RSPAN and the steps for configuring it.

# Remote SPAN (RSPAN)

RSPAN supports source ports, source VLANs, and destinations on different switches, facilitating remote monitoring of multiple switches across networks. RSPAN uses a Layer 2 VLAN to carry SPAN traffic between switches.

RSPAN consists of an RSPAN source session, an RSPAN VLAN, and an RSPAN destination session. RSPAN has a source session and a destination session on each switch. To configure an RSPAN source session on one switch, you associate a set of source ports or VLANs with an RSPAN VLAN. To configure an RSPAN destination session on another switch, you associate the destinations with the RSPAN VLAN.

The traffic for each RSPAN session is carried as Layer 2 nonroutable traffic over a user-specified RSPAN VLAN that is dedicated for that RSPAN session in all participating switches. All participating switches must be connected by trunk links. RSPAN source sessions do not copy locally sourced RSPAN VLAN traffic from source trunk ports that carry RSPAN VLANs. Each RSPAN source session can have either ports or VLANs as sources, but not both. The RSPAN source session copies traffic from the source ports or source VLANs and switches the traffic over the RSPAN VLAN to the RSPAN destination session. The RSPAN destination session switches the traffic to the destinations.

Example 30.12 shows the configuration and verification of RSPAN. VLAN 200 is created as the Remote SPAN VLAN for the RSPAN session. Notice that it is created on SW1 and on SW2 as they will both be participating in RSPAN. VLAN 200 is configured as an RSPAN VLAN using the **remote-span** command. The session is configured to monitor GigabitEthernet 0/0, and then the destination is configured with the RSPAN VLAN 200. On SW2, where the RSPAN destination port will be, the source is configured as an RSPAN VLAN, and the destination port is the analyzer port at GigabitEthernet 0/1. The output of the **show monitor session 2** command indicates that this is a remote SPAN session. It also shows the source RSPAN VLAN and destination port for the session.

EXAMPLE 30.12 **Configuring and Verifying RSPAN**

```
SW1#
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# vlan 200
SW1(config-vlan)# name REMOTE-SPAN-VLAN
SW1(config-vlan)# remote-span
```

```
SW1(config-vlan)# exit
SW1(config)# monitor session 2 source interface gigabitEthernet 0/0
SW1(config)# monitor session 2 destination remote vlan 200
SW1(config)# exit
SW1#
SW1# show monitor session 2
Session 2
---------
Type                      : Remote Source Session
Source Ports              :
    Both                  : Gi0/0
Dest RSPAN VLAN           : 200
SW1#

SW2#
SW2# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW2(config)# vlan 200
SW2(config-vlan)# name REMOTE-SPAN-VLAN
SW2(config-vlan)# remote-span
SW2(config-vlan)# exit
SW2(config)# monitor session 2 source remote vlan 200
SW2(config)# monitor session 2 destination interface
gigabitEthernet 0/1
SW2(config)# exit
SW2#
SW2# show monitor session 2
Session 2
---------
Type                    : Remote Destination Session
Source RSPAN VLAN       : 200
Destination Ports       : Gi0/1
    Encapsulation       : Native
SW2#
```

**ExamAlert**

Before taking the ENCOR exam, make sure you understand the use case for ERSPAN and the steps involved in configuring it.

# Encapsulated Remote SPAN (ERSPAN)

ERSPAN supports source ports, source VLANs, and destinations on different switches across Layer 3 links, providing remote monitoring of multiple

switches across a network. ERSPAN uses a GRE tunnel to carry traffic between switches. ERSPAN consists of an ERSPAN source session, routable ERSPAN GRE-encapsulated traffic, and an ERSPAN destination session. You separately configure ERSPAN source sessions and destination sessions on different switches.

To configure an ERSPAN source session on one switch, you associate a set of source ports or VLANs with a destination IP address, ERSPAN ID number, and, optionally, a VRF instance name. To configure an ERSPAN destination session on another switch, you associate the destination with the source IP address, ERSPAN ID number, and, optionally, a VRF instance name.

ERSPAN source sessions do not copy locally sourced RSPAN VLAN traffic from source trunk ports that carry RSPAN VLANs. ERSPAN source sessions do not copy locally sourced ERSPAN GRE-encapsulated traffic from source ports. Each ERSPAN source session can have either ports or VLANs as sources, but not both.

An ERSPAN source session copies traffic from the source ports or source VLANs and forwards the traffic using routable GRE-encapsulated packets to the ERSPAN destination session. The ERSPAN destination session switches the traffic to the destinations.

Finally, you can configure local SPAN sessions, RSPAN source sessions, and ERSPAN source sessions to monitor the following traffic:

- **Ingress traffic:** In the ingress SPAN process, you copy traffic received by the sources (that is, ingress traffic).
- **Egress traffic:** In the egress SPAN process, you copy traffic transmitted from the sources (that is, egress traffic).

Example 30.13 shows the configuration and verification of ERSPAN. The session type is configured as the source on SW3. The monitor source is interface GigabitEthernet 0/0 for Rx traffic. To configure the destination on SW3, you need to enter the destination configuration mode by using the **destination** command. The destination IP address 10.10.20.1 is specified. The unique identifier for the destination session is specified with the **erspan-id** command. The source IP address or origin of the ERSPAN session is 10.10.10.1. This example specifies a TTL value of 32. On SW4, the destination session is configured to send ERSPAN ID2 traffic to the interface GigabitEthernet 0/1. Verification is done using the **show monitor session erspan-source session** command.

EXAMPLE 30.13 **Configuring and Verifying ERSPAN**

```
SW3#
SW3# configure terminal
SW3(config)# monitor session 1 type erspan-source
SW3(config-mon-erspan-src)# description Server1 traffic
SW3(config-mon-erspan-src)# source interface GigabitEthernet 0/0 rx
SW3(config-mon-erspan-src)# filter vlan 10
SW3(config-mon-erspan-src)# no shutdown
SW3(config-mon-erspan-src)# destination
SW3(config-mon-erspan-src-dst)# ip address 10.10.20.1
SW3(config-mon-erspan-src-dst)# erspan-id 2
SW3(config-mon-erspan-src-dst)# origin ip address 10.10.10.1
SW3(config-mon-erspan-src-dst)# ip ttl 32
SW3(config-mon-erspan-src)# exit
SW3(config)# end
SW3#

SW4#
SW4# configure terminal
SW4(config)# monitor session 1 type erspan-destination
SW4(config-erspan-dst)# destination interface GigabitEthernet 0/1
SW4(config-erspan-dst)# source
SW4(config-erspan-dst-src)# ip address 10.10.20.1
SW4(config-erspan-dst-src)# erspan-id 2
SW4(config-erspan-dst-src)# end
SW4#

SW3# show monitor session erspan-source session

Type : ERSPAN Source Session
Status : Admin Enabled
Source Ports :
RX Only : Gi0/0
Destination IP Address : 10.10.20.1
Destination ERSPAN ID : 2
Origin IP Address : 10.10.10.1
IPv6 Flow Label : None
```

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which version of SPAN requires the source and destination of a session to be on the same device?
   - ❍ **A.** Local SPAN
   - ❍ **B.** RSPAN
   - ❍ **C.** ERSPAN
   - ❍ **D.** GRE

2. Which of the following can be used for capturing packets from one device and sending the capture across a Layer 3 routed link to another destination?
   - ❍ **A.** SPAN
   - ❍ **B.** RSPAN
   - ❍ **C.** ERSPAN
   - ❍ **D.** NetFlow

3. Which Cisco IOS feature allows for the monitoring of traffic on one or more ports or VLANs and sends the traffic to one or more destinations?
   - ❍ **A.** SPAN
   - ❍ **B.** RSPAN
   - ❍ **C.** GRE
   - ❍ **D.** ERSPAN

## Answers

1. **A** is correct. A local SPAN session can only be configured on a local switch, and the source and destination of the session must be on the same device.
2. **C** is correct. The ERSPAN source session copies traffic from the source ports or source VLANs and forwards the traffic using routable GRE-encapsulated packets to the ERSPAN destination session.
3. **D** is correct. ERSPAN provides the flexibility to monitor traffic on one or more ports or VLANs and send the traffic to one or more destinations.

# Review Questions

1. What element of a syslog message refers to the sources and cause of a system message?
   - ❍ **A.** Sequence number
   - ❍ **B.** Timestamp
   - ❍ **C.** Severity
   - ❍ **D.** Facility

2. When logging severity level 2 is configured, what is actually logged? (Choose all that apply.)
   - ❍ **A.** Emergency
   - ❍ **B.** Notification
   - ❍ **C.** Alert
   - ❍ **D.** Critical

3. Which of the following are components for configuring Flexible NetFlow? (Choose four.)
   - ❍ **A.** Flow record
   - ❍ **B.** Flow monitor
   - ❍ **C.** Flow exporter
   - ❍ **D.** Sequence number
   - ❍ **E.** Flow sampler

4. What type of SPAN requires a special VLAN for moving the monitored traffic?
   - ❍ **A.** SPAN
   - ❍ **B.** RSPAN
   - ❍ **C.** ERSPAN
   - ❍ **D.** GRE

5. What command is used to show the type of session, the source port for each traffic direction, and the destination port for SPAN sessions?
   - ❍ **A.** **show session**
   - ❍ **B.** **show logging**
   - ❍ **C.** **show monitor session**
   - ❍ **D.** **show flow monitor**

# Answers to Review Questions

1. **D** is correct. A facility indicates the sources and cause of a system message. A facility can be a hardware device, a protocol, or a module of the system software.
2. **A**, **C**, and **D** are correct. Logging severity level 2 (critical) includes severity levels 1 (alert) and 0 (emergency) as well.
3. **A**, **B**, **C**, and **E** are correct. Flow monitor, flow record, flow exporter, and flow sampler are all components for configuring Flexible NetFlow.
4. **B** is correct. The traffic for each RSPAN session is carried as Layer 2 nonroutable traffic over a user-specified RSPAN VLAN dedicated to that RSPAN session in all participating switches.
5. **C** is correct. **show monitor session** is used to show the type of session, the source port for each traffic direction, and the destination port for SPAN sessions.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers IP SLA and DNA Center.

CHAPTER 31

# IP SLA and DNA Center

**This chapter covers the following official ENCOR 350-401 exam objectives:**

- 4.5 Configure and verify IPSLA
- 4.6 Describe DNA Center workflows to apply network configuration, monitoring, and management

This chapter looks at two separate topics of the network assurance portion of the ENCOR 350-401 exam. It starts by looking at how to use IP Service Level Agreement (SLA) to actively monitor and report on the network performance between multiple network locations or across multiple network paths. It then looks at the use cases for IP SLA, where the measurements provided can be used for troubleshooting, problem analysis, and design of network topologies. This chapter also examines the configuration and verification of IP SLA. The second part of this chapter examines concepts related to Cisco Digital Network Architecture (DNA) Center, with a focus on the Assurance section and associated workflows for troubleshooting and diagnostics.

**This chapter covers the following technology topics:**

- IP SLA Overview
- Cisco DNA Center Assurance

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. What is the ICMP echo operation in IP SLA configuration used for?
2. Which DNA Center component translates business intent into network policies and applies those policies, such as access controls, traffic routing, and quality of service, consistently over the entire wired and wireless infrastructure?

### Answers

1. The ICMP echo operation uses IP probes to measure the end-to-end response time between a Cisco IOS device and any other device. In addition, it can be used to test/monitor basic IP connectivity between a Cisco IOS device and any other device.
2. Policy

# IP SLA Overview

### ExamAlert

Before taking the ENCOR exam, make sure you completely understand the use of IP SLAs. You need to know that an IP SLA Responder is not a mandatory component in an IP SLA setup, but it can provide accurate measurements without the need for dedicated probes.

Network connectivity from the enterprise campus to the WAN and Internet, as well as connectivity to branches and data centers, has become increasingly critical for companies. Downtime or degradation in service can adversely affect revenue. There needs to be some form of predictability with IP services when a company is connected to networks in various forms. A service-level agreement (SLA) is a contract between a service provider and a customer in which some form of guarantee is provided to the customer about the level of user experience. An SLA typically outlines the minimum level of service and the expected level of network service and performance that a company should expect from the service provider.

Internet Protocol Service Level Agreement (IP SLA) is part of the Cisco IOS software that actively monitors and reports on network performance. It does this by generating and actively monitoring traffic continuously across the network. IP SLA tests can use the following operations:

- ICMP
- HTTP
- FTP
- VOIP
- DHCP
- DNS

A Cisco IOS router running IP SLA can monitor and report on traffic in real time. The technical components of an SLA contain a guaranteed level for network availability and network performance in terms of the round-trip time (RTT). An SLA may measure network response time in terms of latency, jitter, and packet loss. IP SLA can be configured to report on these various statistics:

- Connectivity
- Jitter
- Response time
- Packet loss
- Server/website responses and downtime
- Delay
- Voice quality score

Let us look at some of the benefits of using IP SLA. This is not an exhaustive list, but these are the most common benefits that drive a network administrator to implement IP SLA:

- Service-level agreement monitoring, measurement, and verification
- Network performance monitoring, including the following:
  - Measurement of jitter, latency, or packet loss in the network
  - Provision of continuous, reliable, and predictable measurements
- IP service network health assessment to verify that the existing QoS is adequate for new IP services

- Edge-to-edge network availability monitoring for proactive verification and connectivity testing of network resources (for example, showing the network availability of an FTP server used to store business-critical data from a remote site)
- Troubleshooting of network operation by providing consistent, reliable measurement that immediately identifies problems and saves troubleshooting time

Before getting into the configuration of IP SLA, let us look at what is necessary and what is optional to make IP SLA work. First, you need an IP SLA router, which generates the traffic, and an IP SLA responder. An IP SLA responder is not required for IP SLA to function, but using it does enable detailed information gathering and reporting. You cannot set up an IP SLA responder on non-Cisco devices, and Cisco IOS IP SLA can send operational packets only to services native to those devices. Therefore, the source that is generating the traffic toward the responder needs to be a Cisco IOS device.

In its basic form, the IP SLA responder is a component embedded in the destination Cisco IOS device that allows the system to anticipate and respond to IP SLA request packets. An IP SLA responder can provide accurate measurements without the need for dedicated probes. You do not need to enable the responder on the destination device for all IP SLA operations. For example, a responder is not required for services that are already provided by the destination router (such as Telnet, ICMP, or HTTP).

This section examines how to analyze the IP service level by looking at the configuration and verification of three common use cases for IP SLA: ICMP echo operation, monitoring of HTTP destinations, and measurement of IP SLA UDP jitter operation.

The ICMP echo operation uses IP to measure end-to-end response time between a Cisco IOS device and any other device. Response time is computed by measuring the time between sending an ICMP echo request message to the destination and receiving an ICMP echo reply. It is common to use IP SLAs ICMP-based operations or **ping**-based dedicated probes for response time measurements between the source IP SLA device and the destination IP device. The IP SLA ICMP echo operation follows the same specifications as ICMP **ping** testing, as the two methods result in the same response times.

Figure 31.1 shows how **ping** is used with an ICMP echo operation to measure response time between the source IP SLA device and the destination IP device.

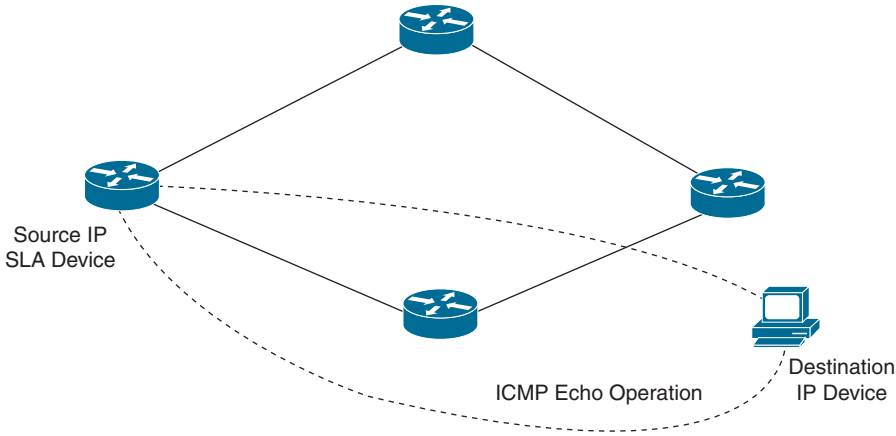



FIGURE 31.1 **ICMP Echo Operation**

There are many optional parameters available for configuring an IP SLA operation. However, the examples in this chapter focus on the basic setup and scheduling of IP SLA ICMP echo operations. The following steps are involved:

1. To configure an IP SLA operation, use the **ip sla** *operation-number* command to enter the IP SLA configuration mode (where *operation-number* is the configuration for the individual IP SLA probe). This is necessary to configure multiple IP SLA instances on a single device, where they are all doing different operations or verification tasks. Once you are in the IP SLA configuration mode, you use the command **icmp-echo** {*destination-ip-address* | *destination-hostname*} [**source-ip** {*ip-address* | *hostname*} | **source-interface** *interface-name*] to configure the destination IP address of the device or host to be monitored.
2. Specify how often the ICMP echo operation should run by using the **frequency** *seconds* command.
3. When the IP SLA configuration is complete, schedule and activate the IP SLA operation that has been configured by using the **ip sla schedule** *operation-number* [**life** {**forever** | *seconds*}] [**start-time** {[*hh:mm:ss*] [*month day* | *day month*] | **pending** | **now** | **after** *hh:mm:ss*}] [**ageout** *seconds*] [**recurring**] command.
4. For verification, use the command **show ip sla configuration.**

Example 31.1 shows a simple configuration and verification of ICMP echo operation. This example specifies the IP SLA operation number by using the command **ip sla 1** command. This example shows an ICMP echo operation to destination 10.10.10.99 using Loopback0 as the source interface. The ICMP

echo operation is specified to run every 300 seconds. This example also configures a schedule to run forever and to run immediately, via the **ip sla schedule 1 life forever start-time now** command. Verification is done using the **show ip sla configuration 1** command, whose output shows the type of operation as **icmp-echo**, the destination address, and the frequency.

EXAMPLE 31.1 **Configuring and Verifying ICMP Echo Operation**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# ip sla 1
R1(config-ip-sla)# icmp-echo 10.10.10.99 source-interface Loopback0
R1(config-ip-sla-echo)# frequency 300
R1(config-ip-sla-echo)# exit
R1(config)# ip sla schedule 1 life forever start-time now
R1(config)# exit
R1#
R1# show ip sla configuration 1
IP SLAs Infrastructure Engine-III
Entry number: 1
Owner:
Tag:
Operation timeout (milliseconds): 5000
Type of operation to perform: icmp-echo
Target address/Source interface: 10.10.10.99/Loopback0
Type Of Service parameter: 0x0
Request size (ARR data portion): 28
Data pattern: 0xABCDABCD
Verify data: No
Vrf Name:
Schedule:
   Operation frequency (seconds): 60  (not considered if randomly
scheduled)
   Next Scheduled Start Time: Start Time already passed
   Group Scheduled : FALSE
   Randomly Scheduled : FALSE
   Life (seconds): Forever
   Entry Ageout (seconds): never
   Recurring (Starting Everyday): FALSE
   Status of entry (SNMP RowStatus): Active
Threshold (milliseconds): 5000
Distribution Statistics:
   Number of statistic hours kept: 2
   Number of statistic distribution buckets kept: 1
   Statistic distribution interval (milliseconds): 20
Enhanced History:
History Statistics:
```

```
   Number of history Lives kept: 0
   Number of history Buckets kept: 15
   History Filter Type: None

R1# n
```

Let us now look at another example of using IP SLA to monitor HTTP destinations. This can be done by using the HTTP GET operation of IP SLA, which involves the following steps:

1. To configure IP SLA to monitor HTTP destinations for operations, use the **ip sla** *operation-number* command to enter the IP SLA configuration mode. Then configure the HTTP GET probe by issuing the command **http** {**get** | **raw**} *url* [**name-server** *ip-address*] [**version** *version-number*] [**source-ip** {*ip-address* | *hostname*}] [**source-port** *port-number*] [**cache** {**enable** | **disable**}] [**proxy** *proxy-url*].
2. Specify how often the ICMP echo operation should run by using the **frequency** *seconds* command.
3. When the IP SLA configuration is complete, schedule and activate the IP SLA operation by using the command **ip sla schedule** *operation-number* [**life** {**forever** | *seconds*}] [**start-time** {[*hh:mm:ss*] [*month day* | *day month*] | **pending** | **now** | **after** *hh:mm:ss*}] [**ageout** *seconds*] [**recurring**].
4. For verification, use the command **show ip sla configuration.**

Example 31.2 shows a simple configuration and verification of the IP SLA HTTP GET operation. It configures the HTTP GET probe to monitor the destination at 10.10.10.99. The HTTP GET operation is specified to run every 90 seconds. This example also configures a schedule to run forever and to run immediately. Verification is done using the **show ip sla configuration 2** command. This shows the type of operation as HTTP, the destination address, and the frequency.

EXAMPLE 31.2 **Configuring and Verifying IP HTTP GET Operation**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# ip sla 2
R1(config-ip-sla)# http get http://10.10.10.99
R1(config-ip-sla-http)# frequency 90
R1(config-ip-sla-http)# exit
R1(config)# ip sla schedule 2 start-time now life forever
R1(config)# exit
R1#
```

```
R1# show ip sla configuration 2
IP SLAs Infrastructure Engine-III
Entry number: 2
Type of operation to perform: http
Target address/Source address: 10.10.10.99/0.0.0.0
Target port/Source port: 80/0
Type Of Service parameters: 0x0
Vrf Name:
HTTP Operation: get
HTTP Server Version: 1.0
URL: http://10.10.10.99
Proxy:
Raw String(s):
Cache Control: enable
Owner:
Tag:
Operation timeout (milliseconds): 60000
Schedule:
   Operation frequency (seconds): 90  (not considered if randomly
scheduled)
   Next Scheduled Start Time: Start Time already passed
   Group Scheduled : FALSE
   Randomly Scheduled : FALSE
   Life (seconds): Forever
   Entry Ageout (seconds): never
   Recurring (Starting Everyday): FALSE
   Status of entry (SNMP RowStatus): Active
Threshold (milliseconds): 5000
Distribution Statistics:
   Number of statistic hours kept: 2
   Number of statistic distribution buckets kept: 1
   Statistic distribution interval (milliseconds): 20
History Statistics:
   Number of history Lives kept: 0
   Number of history Buckets kept: 15
   History Filter Type: None

R1#
```

Example 31.3 shows the configuration and verification of multiple IP SLA instances to measure both ICMP echo and UDP jitter. There could be multiple IP SLA instances on the same device, with each instance doing a different operation or verification task. For this example, assume that R1 is an HQ router connecting to R2, which is a branch router. Say that you are connecting over an MPLS cloud, and you are going to use SLA 1 to monitor ICMP echo operation and SLA 2 to monitor UDP jitter. R2 will serve as the responder. The UDP jitter test is ideal if you want to make sure the ISP is providing the level of service agreed on (especially if you are running time-sensitive traffic like voice).

The focus of this example is to demonstrate the creation of multiple SLAs and a UDP jitter test (SLA2). (For details on ICMP echo operation, see Example 30.2, earlier in this section.) IP SLA2 is configured for destination 10.10.10.2 with UDP port set to 65060. Every 20 seconds, R1 will transmit 20 160-bytes packets sent 15 milliseconds apart. On R2, which is simulating the branch, only the **ip sla responder** command is needed.

EXAMPLE 31.3 **Configuring ICMP Echo and UDP Jitter Operation**

```
R1#
R1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)# ip sla 1
R1(config-ip-sla)# icmp-echo 10.10.10.2
R1(config-ip-sla-echo)# ip sla schedule 1 life forever start-time now
R1(config)# ip sla 2
R1(config-ip-sla)# udp-jitter 10.10.10.2 65060 num-packets 20 interval
15
R1(config-ip-sla-jitter)# request-data-size 160
R1(config-ip-sla-jitter)# frequency 20
R1(config-ip-sla-jitter)# ip sla schedule 2 start-time now
R1(config)#

R2#
R2# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R2(config)# ip sla responder
R2(config)#
```

Example 30.4 shows the verification of the ICMP echo and UDP jitter operation configuration from Example 30.3. The **show ip summary** and **show ip statistics** command output shows that both SLAs are reporting OK status, and the UDP jitter SLA is gathering latency and jitter times between R1 and R2.

EXAMPLE 30.4 **Verifying ICMP Echo and Jitter Operation**

```
R1# show ip sla summary
IPSLAs Latest Operation Summary
Codes: * active, ^ inactive, ~ pending

ID           Type        Destination       Stats       Return
Last
                                           (ms)        Code        Run
-------------------------------------------------------------------------
--
```

```
*1            icmp-echo   10.10.10.2         RTT=2        OK            4
seconds ago



*2            udp-jitter  10.10.10.2         RTT=3        OK            2
seconds ago



R1# show ip sla statistics
IPSLAs Latest Operation Statistics

IPSLA operation id: 1
        Latest RTT: 2 milliseconds
Latest operation start time: 01:02:59 UTC Sun Nov 28 2021
Latest operation return code: OK
Number of successes: 1
Number of failures: 1
Operation time to live: Forever



IPSLA operation id: 2
Type of operation: udp-jitter
        Latest RTT: 3 milliseconds
Latest operation start time: 01:03:01 UTC Sun Nov 28 2021
Latest operation return code: OK
RTT Values:
        Number Of RTT: 20               RTT Min/Avg/Max: 2/3/5
milliseconds
Latency one-way time:
        Number of Latency one-way Samples: 0
        Source to Destination Latency one way Min/Avg/Max: 0/0/0
milliseconds
        Destination to Source Latency one way Min/Avg/Max: 0/0/0
milliseconds
Jitter Time:
        Number of SD Jitter Samples: 19
        Number of DS Jitter Samples: 19
        Source to Destination Jitter Min/Avg/Max: 0/1/2 milliseconds
        Destination to Source Jitter Min/Avg/Max: 0/2/3 milliseconds
Over Threshold:
        Number Of RTT Over Threshold: 0 (0%)
Packet Loss Values:
        Loss Source to Destination: 0
        Source to Destination Loss Periods Number: 0
```

```
        Source to Destination Loss Period Length Min/Max: 0/0
        Source to Destination Inter Loss Period Length Min/Max: 0/0
        Loss Destination to Source: 0
        Destination to Source Loss Periods Number: 0
        Destination to Source Loss Period Length Min/Max: 0/0
        Destination to Source Inter Loss Period Length Min/Max: 0/0
        Out Of Sequence: 0      Tail Drop: 0
        Packet Late Arrival: 0  Packet Skipped: 0
Voice Score Values:
        Calculated Planning Impairment Factor (ICPIF): 0
        Mean Opinion Score (MOS): 0
Number of successes: 4
Number of failures: 0
Operation time to live: 3530 sec
R1#
```

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. True or false: An IP SLA responder is required for IP SLA to work.
   - ❍ **A.** True
   - ❍ **B.** False

2. True or false: When configuring IP SLA, you cannot configure multiple IP SLA instances on a single device.
   - ❍ **A.** True
   - ❍ **B.** False

## Answers

1. **B** is correct. An IP SLA responder is not required for IP SLA to work, but using it does enable detailed information gathering and reporting.
2. **B** is correct. You can configure multiple IP SLA instances on a single device, where they are all doing different operations or verification tasks.

# Cisco DNA Center Assurance

**ExamAlert**

Before taking the ENCOR exam, make sure you understand the advantages of using the Assurance section of DNA Center for monitoring your network's health. Also make sure you understand the basics of the Client Health dashboard and how scores are used.

This section provides a brief overview of Cisco DNA Center and looks at the four general sections of the DNA Center dashboard: Design, Policy, Provision, and Assurance. This section focuses on the Assurance section, as this is the primary focus of the network assurance portion of the ENCOR 350-401 exam objectives.

Cisco DNA Center is a management and control platform that simplifies and streamlines network operations in a software-defined networking (SDN) environment. This extensible software platform includes integrated tools for NetOps, SecOps, DevOps, and the Internet of Things (IoT). Without Cisco DNA Center, such complete functionality would be possible only with the purchase and operation of multiple third-party software tools. These are some of the advantages of having all the core network tools integrated into the Cisco DNA Center software platform:

- Multiple tools with multiple interfaces add complexity, which increases the possibility of errors in configuration and management. This can be especially harmful when errors in security settings lead to open vulnerabilities.
- Changing between program interfaces during network operations is time-consuming and can make even simple changes or troubleshooting tasks take much longer.
- Third-party platforms cannot support the same levels of device management and control as those that are integrated and designed to work together.
- Automatic troubleshooting with guided remediation is extremely complex in today's virtualized networks. Third-party tools can often tell you if a problem is due to the network or caused by an application, but they can't offer guided remediation without true integration between the tools that control virtualization, analytics, and automation.

- Real intent-based networking requires extensive real-time data flow between the operational tools that are core to the network. The management of network configuration, security, analytics, and automation comes together to deliver the true business intent of the operation.
- Cisco DNA Center is an open and extensible platform that allows third-party applications and processes to exchange data and intelligence with your network. This improves IT operations by automating workflow processes based on network intelligence coming from Cisco DNA Center.

Cisco DNA Center offers a single dashboard you can use for every core function in a network. Cisco DNA Center is the network management system, the foundational controller, and the analytics platform at the heart of Cisco's intent-based networking. Beyond device management and configuration, DNA Center is a set of software solutions that provide the following:

- A management platform for the entire network
- An intent-based networking controller for automation of policies, segmentation, and services configurations
- An assurance engine to guarantee the best network experience for all users

Cisco DNA Center software sits on the Cisco DNA Center appliance and controls all of your Cisco devices; it supports fabric and nonfabric deployments. From its main menu, Cisco DNA Center has four general sections aligned to IT workflows:

- **Design:** This section enables you to design a network for consistent configuration by device and by site. Physical maps and logical topologies help provide quick visual reference. The direct import option brings in existing maps, images, and topologies directly from the Cisco Prime Infrastructure and the Cisco Application Policy Infrastructure Controller Enterprise Module (APIC-EM), which makes upgrades easy and quick. Device configurations by site can be consolidated in a "golden image" that can be used to provision new network devices automatically. You can either prestage these new devices by associating the device details and mapping to a site, or you can claim them upon connection and map them to the site.

- **Policy:** This section translates business intent into network policies and applies those policies—such as access control, traffic routing, and quality of service—consistently over entire wired and wireless infrastructures. Policy-based access control and network segmentation are critical functions of the SD-Access solution built from the Cisco DNA Center and the Cisco Identity Services Engine (ISE).

  Cisco AI Network Analytics and the Cisco Group-Based Policy Analytics running in Cisco DNA Center identify endpoints, group similar endpoints, and determine group communication behavior. Cisco DNA Center facilitates the creation of policies that determine the form of communication allowed between groups and between members of each group. ISE activates the underlying infrastructure and segments the network, creating a virtual overlay to follow these policies consistently. Such segmenting implements zero-trust security in the network environment and reduces risk, contains threats, and helps verify regulatory compliance by giving endpoints just the right level of access.

- **Provision:** Once you have created policies in DNA Center, provisioning is a simple drag-and-drop task. Each profile (called a security group tag [SGT]) in the DNA Center inventory list is assigned a policy, and this policy will always follow the identity. New devices added to the network are assigned to an SGT based on identity. The process is completely automated and zero-touch.

- **Assurance:** The Assurance section, which uses artificial intelligence/machine learning (AI/ML), enables every point on a network to become a sensor, sending continuous streaming telemetry on application performance and user connectivity in real time. The dashboard shows network health, and it flags issues. Guided remediation automates resolution to keep the network performing optimally and requiring less mundane troubleshooting work. This provides a consistent experience and proactive optimization of your network, and the network administrator can spend less time troubleshooting tasks.

Figure 31.2 shows the Cisco DNA Center dashboard.

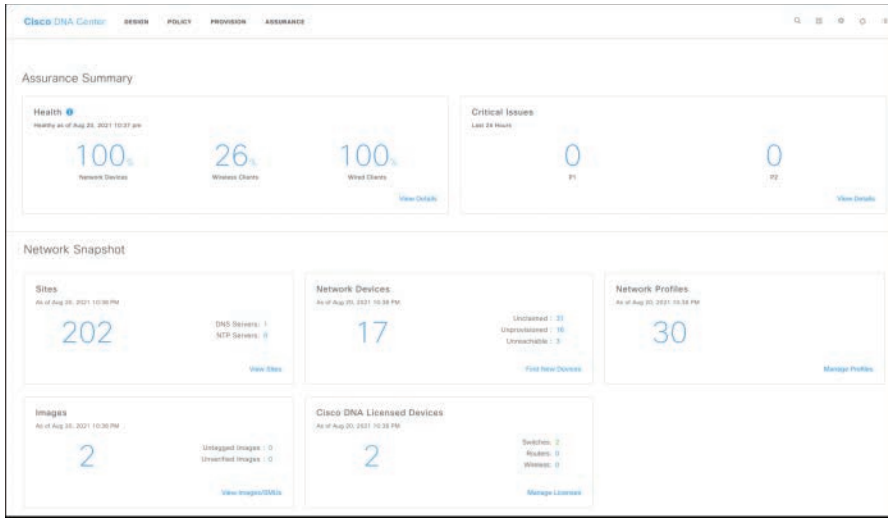



FIGURE 31.2 **Cisco DNA Center Dashboard**

The Assurance section of Cisco DNA Center offers some of the following capabilities (as well as others):

- Cisco SD-Access fabric configuration
- Software image management
- Simplified provisioning for devices
- Wireless network management
- Simplified security policies
- Configuration templates
- Third-party integration
- Network assurance
- Plug and play

The Assurance section of Cisco DNA Center incorporates 30+ years of Cisco Technical Assistance Center (TAC) experience into a tool that uses machine learning to diagnose issues in a network environment. In addition to simply identifying and diagnosing problems, the Assurance section provides guided remediation steps to address the issue. The Assurance section of Cisco DNA Center provides an overview of how the network is performing—from an overall health perspective to a client perspective. It shows the top issues that are

impacting the network environment and provides health scores for each section, enabling you to see, at a glance, how the network is performing.

The Assurance section includes a number of different dashboards and options:

- **Network Health and Client Health:** These dashboards give a quick overview of the health of every wired and wireless network device and clients on the network. They offer a general overview of the operational status of every network device provisioned from Cisco DNA Center. Any devices that are poorly connected are highlighted, and remediation options are suggested.
- **Device 360/Client 360:** This option displays device or client connectivity from any context. It provides a detailed view of the performance of any device or client over time and from any application context. It includes information on topology, throughput, and latency from different times and applications. Using this information, a network administrator can do granular troubleshooting in seconds.
- **Network Time Travel:** This option allows a network administrator to see device or client performance in a timeline view to understand the network state when an issue occurred and enables you to go back in time and see the cause of that network issue instead of trying to re-create the issue in a lab.
- **Path Trace:** This option allows a network administrator to visualize the path of an application or a service from the client through all devices and to the server. By using this option, you can instantly perform a common and critical troubleshooting task that normally requires minutes by simply clicking on a client or an application.
- **Wireless Active Sensor:** This option provides location-based sensor heatmaps to quickly identify failed tests and potential network issues. You can simulate real-world client experiences to validate wireless performance for critical venues and high-value locations such as conference halls and meeting rooms.
- **Wi-Fi 6 Readiness:** This dashboard verifies your hardware and configuration compatibility for the new Wi-Fi standard and locates areas most served by an upgrade. After upgrading, advanced wireless analytics indicate performance and capacity gains as a result of the Wi-Fi 6 deployment. This dashboard improves the visibility of the wireless network by identifying the generation/version of each access point in use as well as each client connected. You can also look at wireless load, throughput, and

performance from many directions. This dashboard allows for upgrading where and when it makes sense to do so and documents the results.

- **iOS and Samsung Client Device Analytics:** This option provides communication from smartphones running Apple iOS or Samsung Android to the Assurance section of Cisco DNA Center. A smartphone client sends error codes and other wireless diagnostic information and allows Cisco DNA Center to provide highly accurate remediation recommendations when an iOS or Samsung mobile client has issues with wireless connectivity to the network.
- **Rogue Management:** This option allows for detection of unauthorized access points plugged into local switches or access points with the same SSID that are not connected as part of the wired network. It allows for increased security and control of wireless networks.
- **User-Defined Networking:** This option allows IT to give end users control of their own wireless network partition on a shared network. End users can then remotely and securely deploy their devices on the network.
- **Machine learning algorithms:** As network conditions change, context-aware baselining captures the relationships between metrics and constantly updates an optimal curve (regression) for performance. Specific issues can be identified when they deviate from this ever-changing baseline. This allows for the updating of the preferred performance curve in real time as network conditions change. Issues raised are based on current and real network conditions rather than on a static model, resulting in fewer issues to troubleshoot.

Figure 31.3 shows the Cisco DNA Center Overall Health dashboard. It shows the network devices and the network score. The network score is the percentage of healthy (good) devices (routers, switches, wireless controllers, and access points) in the overall enterprise. Notice that the categories in the Network Devices section are Router, Core, Distribution, Access, Wireless Controller, and Access Point. The Wired Clients and Wireless Clients sections show the percentage of healthy (good) wired and wireless client devices in the overall enterprise network. The Top 10 Issue Types section shows issues that must be addressed. Issues can show up as either resolved or ignored for a certain number of hours.

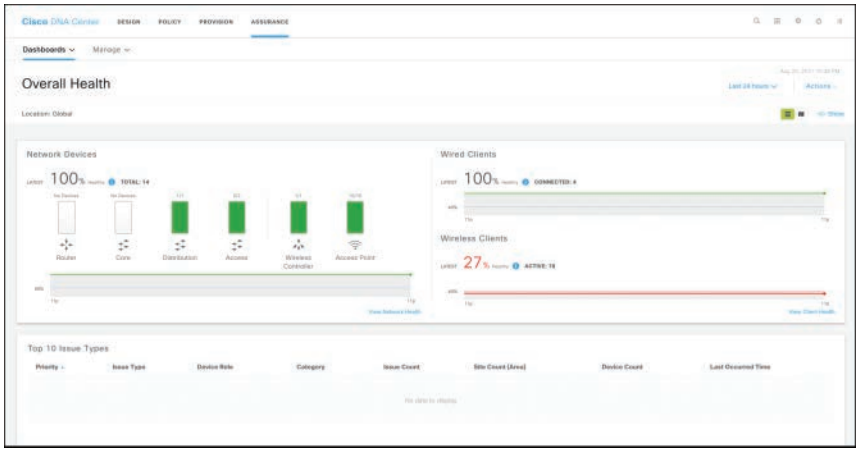



FIGURE 31.3 **Cisco DNA Center Overall Health Dashboard**

Figure 31.4 shows the Cisco DNA Center Client Health dashboard. The client health score is displayed for both wired and wireless clients. It is a percentage of the number of healthy clients (a health score from 8 to 10) in a target category, divided by the total number of client devices in that category. The score is calculated every 5 minutes. The Client Onboarding Times section indicates the experience of the client device while connecting to the network. A successfully connected device is scored 4; a client being unable to connect to the network is given a score of 0. The Connectivity sections indicate the experience of a client device *after* it is connected to the network. The wireless health score is divided into an RSSI-driven connectivity score and an SNR-driven connectivity score.


FIGURE 31.4 **Cisco DNA Center Client Health Dashboard**

## CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which of the following are DNA Center sections that align with the IT workflow? (Choose four.)
   - ❍ **A.** Provision
   - ❍ **B.** Assurance
   - ❍ **C.** Automation
   - ❍ **D.** Policy
   - ❍ **E.** Design

2. Which Cisco DNA Center Assurance section component allows for context-aware baselining, capturing the relationship between metrics as the network condition changes?
   - ❍ **A.** Network Health dashboard and Client Health dashboard
   - ❍ **B.** Device 360/Client 360
   - ❍ **C.** User-defined networking
   - ❍ **D.** Machine learning algorithms

# Answers

1. **A**, **B**, **D**, and **E** are correct. Provision, Assurance, Policy, and Design are the four components of the DNA Center dashboard that align with the IT workflow.
2. **D** is correct. As the network condition changes, Cisco DNA Center Assurance machine learning algorithms allow for context-aware baselining, capturing the relationship between metrics and constantly updating an optimal curve (regression) for performance.

# Review Questions

1. IP SLA can be used to monitor which of the following? (Choose three.)
   - ❍ **A.** Syslog messages
   - ❍ **B.** Packet loss
   - ❍ **C.** Server/website responses and downtime
   - ❍ **D.** Delay

2. Which Cisco DNA Center Assurance section components include information on topology, throughput, and latency from different times and applications?
   - ❍ **A.** Network Health dashboard and Client Health dashboard
   - ❍ **B.** Device 360/Client 360
   - ❍ **C.** User-Defined Networking
   - ❍ **D.** Machine learning algorithms

# Answers to Review Questions

1. **B**, **C**, and **D** are correct. IP SLA can be configured to report on various statistics. It has no interaction with syslog messages.
2. **B** is correct. The Device 360/Client 360 component of the Cisco DNA Center Assurance section includes information on topology, throughput, and latency from different times and applications.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*
- *CCNP and CCIE Enterprise Core & CCNP Advanced Routing Portable Command Guide*

# What's Next?

If you want more practice on this chapter's exam objectives before you move on, remember that you can access all of the CramQuiz questions on the book's web page. The next chapter covers NETCONF and RESTCONF.

CHAPTER 32

# NETCONF and RESTCONF

**This chapter covers the following official ENCOR 350-401 exam objective:**

- 4.7 Configure and verify NETCONF and RESTCONF

This chapter covers configuration and verification of the Network Configuration (NETCONF) and Representation State Transfer Configuration (RESTCONF) protocols. It examines the characteristics of the NETCONF and RESTCONF protocols as well as the configuration and verification of NETCONF and RESTCONF on the CSR1000v platform.

**This chapter covers the following technology topics:**

- NETCONF
- RESTCONF

## CramSaver

If you can correctly answer these CramSaver questions, save time by skimming the Exam Alerts in this section and then completing the CramQuiz at the end of each section and the Review Questions at the end of the chapter. If you are in doubt at all, read everything in this chapter!

1. What protocol does NETCONF use as the transport between network devices?
2. What do CRUD operations represent?

### Answers

1. SSH
2. RESTCONF uses create, read, update, and delete (CRUD) operations on a conceptual datastore.

# NETCONF

### ExamAlert

For the ENCOR exam, make sure you know that NETCONF is standard based and XML encoded, that it uses RPC for communication between client and server, and that it uses SSH to maintain confidentiality, integrity, and authenticity.

Network Configuration (NETCONF) is a standard-based and XML-encoded protocol that provides the transport needed to communicate the YANG-formatted configuration or operational data request from an application that runs on a centralized management platform to a network device. A network administrator can use NETCONF for configuring or requesting operational data from a device. NETCONF uses a simple Remote Procedure Call (RPC)–based mechanism to facilitate communication between a client (centralized management platform script or application) and a server (network device). It uses Secure Shell (SSH) as the transport across network devices so that confidentiality, integrity, and authenticity can be maintained. Table 32.1 describes some of the NETCONF operations.

TABLE 32.1 **Common NETCONF Operations**

| Operation | Description |
|---|---|
| **get** | Retrieves running configuration and device state information. |
| **get-config** | Retrieves all or part of a specified configuration datastore. |
| **edit-config** | Loads all or part of a specified configuration to the specified target configuration datastore. |
| **copy-config** | Creates or replaces an entire configuration datastore with the contents of another complete configuration datastore. The target datastore is replaced, if it exists. |
| **delete-config** | Deletes a configuration datastore. The running configuration datastore cannot be deleted. |

The NETCONF client establishes an SSH connection with the NETCONF server. The NETCONF server then sends a hello message encoded with XML. It declares the NETCONF capabilities that the NETCONF server is capable of, the YANG data models that the device knows, some other proprietary models, and other information. The NETCONF client then replies with the NETCONF client capabilities. When the capabilities between them match, the client can send a series of RPC requests encoded in XML format and send the requests to the server by using SSH. The server then responds with an RPC reply encoded in XML. The request and the response contents are fully described in XML schemas, allowing the client and server to recognize the syntax constraints imposed on the exchange. Depending on the device capabilities, the content data elements may be represented by a native YANG model, an Open Config YANG model, or some other data model.

The key benefit of NETCONF is that it is a standards-based network management protocol. Also, management can be performed both at a single network device level and a network-wide level. The use of NETCONF is considered transaction-based management, in which all configurations are applied or none are applied, thus preventing inconsistent states. NETCONF also provides rich capabilities such as locking, confirmed commit, manipulate, and validate without affecting the running configuration of a network device by using the candidate datastore feature if the device supports that datastore.

NETCONF defines one or more configuration datastores and allows for configuration operations on those datastores. The configuration datastore serves as the complete set of configuration data necessary to get the network device from its initial state to the desired operational state. The configuration datastores include the following:

- **Startup:** The startup configuration datastore holds the configuration loaded by the device when it boots.

- **Running:** The running configuration datastore holds the complete current configuration on the device.
- **Candidate:** The candidate configuration datastore provides a temporary workspace in which a copy of the device's running configuration is stored. It can be manipulated without affecting the device's current configuration.

Now that we have covered the main highlights of NETCONF, let us examine how to configure and verify its setup on a CSR1000v platform. The following high-level steps are necessary to configure and verify NETCONF:

1. Create a user with privilege level 15 to start working with NETCONF APIs.
2. Allow SSH access to the device. When generating the RSA key, ensure that the size of the key modulus is 2048 bits.
3. Enable the NETCONF interface with the **netconf ssh** and **netconf-yang** commands.
4. Use the **netconf-yang feature candidate-datastore** command to enable the candidate datastore.
5. Display the status of the software processes required to support NETCONF-YANG by using the **show platform software yang-management process** command.
6. Verify the NETCONF-YANG sessions by using the **show netconf-yang sessions** command. Use the **show netconf-yang sessions details** command to get detailed information about the sessions.
7. Display statistics about NETCONF-YANG by using the **show netconf-yang statistics** command.
8. Show information about the NETCONF-YANG datastore by using the **show netconf-yang datastores** command.

Example 32.1 shows the configuration and verification of NETCONF on the CSR1000v platform. To start working with NETCONF APIs, you need a user with privilege level 15, and you need to configure SSH access to the device by using the **netconf ssh** command. To enable NETCONF globally, you use the **netconf-yang** command. Notice that after this command is entered, you get a console message saying that the NETCONF-YANG server has been notified to start; this can take up to 90 seconds. For verification, you use the **show platform software yang-management process** command. A few of the

important parts of this command's output are **ncsshd**, **dmiauthd**, **ndbmand**. **ncsshd** is the NETCONF SSH daemon, **dmiauthd** is the data management interface (DMI) authentication daemon, and **ndbmand** is the NETCONF database manager. You get information about NETCONF-YANG statistics by using the **show netconf-yang statistics** command. You can see active sessions by using the **show netconf-yang sessions** command. You find information about the NETCONF-YANG datastores by using the **show netconf-yang datastores** command.

EXAMPLE 32.1 **Configuring and Verifying NETCONF**

```
Router1#
Router1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
Router1(config)# username admin1 privilege 15 password ExamCram#123
Router1(config)# crypto key generate rsa
The name for the keys will be: Router1.test.local
Choose the size of the key modulus in the range of 512 to 4096 for
your
  General Purpose Keys. Choosing a key modulus greater than 512 may
take
  a few minutes.

How many bits in the modulus [1024]: 2048
% Generating 2048 bit RSA keys, keys will be non-exportable...
[OK] (elapsed time was 1 seconds)

Router1(config)#
Router1(config)# ip ssh version 2
Router1(config)# line vty 0 4
Router1(config-line)# login local
Router1(config-line)# transport input ssh
Router1(config-line)# exit
Router1(config)# netconf ssh
Router1(config)# netconf-yang
Router1(config)#
*Aug 29 15:43:50.039: %PSD_MOD-5-DMI_NOTIFY_NETCONF_START: R0/0: psd:
PSD/DMI: netconf-yang server has been notified to start
*Aug 29 15:44:47.047: %NDBMAN-5-ACTIVE: R0/0: ndbmand: All data pro-
viders active.
*Aug 29 15:44:56.012: %DMI-5-NACM_INIT: R0/0: dmiauthd: NACM configu-
ration has been set to its initial configuration.
*Aug 29 15:45:05.675: %DMI-5-SYNC_COMPLETE: R0/0: dmiauthd: The run-
ning configuration has been synchronized to the NETCONF running data
store.
```

```
Router1#! for verification:
Router1# show platform software yang-management process
confd            : Running
nesd             : Running
syncfd           : Running
ncsshd           : Running
dmiauthd         : Running
nginx            : Running
ndbmand          : Running
pubd             : Running

Router1# show netconf-yang statistics
netconf-start-time  : 2021-08-29T15:45:06+00:00
in-rpcs             : 0
in-bad-rpcs         : 0
out-rpc-errors      : 0
out-notifications   : 0
in-sessions         : 0
dropped-sessions    : 0
in-bad-hellos       : 0

Router1#! there are not any sessions currently
Router1# show netconf-yang sessions
There are no active sessions

Router1# show netconf-yang datastores
Datastore Name             : running
```

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. Which NETCONF operation loads all or part of a specified configuration to the specified target configuration datastore?
   - ❍ **A. get**
   - ❍ **B. get-config**
   - ❍ **C. edit-config**
   - ❍ **D. copy-config**

2. True or false: The candidate configuration datastore provides a temporary workspace for storing a device's running configuration.

   ❍ **A.** True

   ❍ **B.** False

# Answers

1. **C** is correct. The **edit-config** operation loads all or part of a specified configuration to the specified target configuration datastore.
2. **A** is correct. The candidate configuration datastore provides a temporary workspace in which a copy of a device's running configuration is stored.

# RESTCONF

**ExamAlert**

Before taking the ENCOR exam, make sure you understand what the RESTCONF model is based on and how it differs from NETCONF.

Representational State Transfer Configuration (RESTCONF) is a protocol that provides a programmatic interface based on standard mechanisms used for accessing configuration data, state data, data-model-specific RPC operations, and events, as defined in the YANG model. RESTCONF is a specification that is similar to the NETCONF-YANG interface model. Whereas NETCONF-YANG uses an SSH model, RESTCONF is based on JSON and HTTP. RESTCONF provides a mechanism for accessing data stored in NETCONF datastores over HTTP. The **nginx** process runs if the command **ip http secure-server** or **ip http server** is configured. While the **nginx** process does not need to be in a running state for NETCONF to function properly, it is required for RESTCONF. NGINX is an internal web server that acts as a proxy web server. It provides TLS-based HTTPS. A RESTCONF request that is sent via HTTPS is first received by the NGINX proxy web server, and the request is then transferred to the **confd** web server for further syntax/semantics checking.

RESTCONF helps support a common REST-based programming model for network programming. Using HTTP, RESTCONF combines HTTP's simplicity and the predictability and automation potential of a schema-driven API. RESTCONF uses structured data (JSON) and YANG to provide REST-like APIs, and it enables you to programmatically access different network devices. The RESTCONF APIs use HTTP methods. RESTCONF is a stateless protocol that uses secure HTTP methods to provide create, read, update, and delete (CRUD) operations on a conceptual datastore containing YANG-defined data. Table 32.1 maps CRUD to HTTP methods with RESTCONF.

TABLE 32.1 **CRUD Mapping with RESTCONF**

| Options | Supported Methods |
|---|---|
| GET | Read |
| PATCH | Update |
| PUT | Create or Replace |
| POST | Create or Operations (reload, default) |

| Options | Supported Methods |
|---|---|
| DELETE | Delete the targeted resource |
| HEAD | Header metadata (no response body) |

The following steps show how to configure and verify RESTCONF on the same CSR1000v used in the last example. Steps, like user creation, are skipped in this case:

1. Enable the RESTCONF interface by using the **restconf** command.
2. Enable the secure HTTP (HTTPS) server by using the **ip http secure-server** command.
3. Enable local authentication by using the **ip http authentication** *local* command.
4. When the device boots up with its startup configuration, to verify that the **nginx** process is running, use the **show platform software yang-management process** command.
5. Verify the NETCONF/RESTCONF sessions by using the **show netconf-yang sessions** command. The **show netconf-yang sessions details** command shows detailed information about the sessions.

Example 32.2 shows the configuration and verification of RESTCONF on the CSR1000v platform. You enable the RESTCONF interface by using the **restconf** command. You enable secure HTTP (HTTPS) by using the **ip http secure-server** command. For verification, notice that the **nginx** process is running. Also notice that the **confd** process is running. The NGINX proxy web server transfers requests to the **confd** web server for further syntax/semantics checking.

EXAMPLE 32.2 **Configuring and Verifying RESTCONF**

```
Router1#
Router1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
Router1(config)#restconf
Router1(config)#
*Aug 29 20:57:22.588: %PSD_MOD-5-DMI_NOTIFY_RESTCONF_START: R0/0: psd:
PSD/DMI: restconf server has been notified to start
Router1(config)# ip http secure-server
Router1(config)#
*Aug 29 20:58:10.918: %DMI-5-SYNC_NEEDED: R0/0: dmiauthd: Configura-
tion change requiring running configuration sync detected - 'ip http
secure-server '. The running configuration will be synchronized  to
the NETCONF running data store.
```

```
*Aug 29 20:58:14.557: %DMI-5-SYNC_COMPLETE: R0/0: dmiauthd: The run-
ning configuration has been synchronized to the NETCONF running data
store.
Router1(config)# ip http authentication local
Router1(config)# exit
Router1#
Router1#! for verification:
Router1# show platform software yang-management process
confd            : Running
nesd             : Running
syncfd           : Running
ncsshd           : Running
dmiauthd         : Running
nginx            : Running
ndbmand          : Running
pubd             : Running
Router1#
```

# CramQuiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading this section again until you can.

1. True or false: When a NETCONF client establishes a connection to a NETCONF server, it establishes an SSH connection.
   - ❍ **A.** True
   - ❍ **B.** False

2. What command is used for verifying that the processes for NETCONF and RESTCONF are running after configuration?
   - ❍ **A. show platform software yang-management process**
   - ❍ **B. show netconf-yang sessions**
   - ❍ **C. show netconf-yang sessions details**
   - ❍ **D. show netconf-yang statistics**

# Answers

1. **A** is correct. When a NETCONF client (that is, central management device) establishes a connection to a NETCONF server (that is, network device), it establishes an SSH connection.
2. **A** is correct. You can display the status of the software processes required to support NETCONF and RESTCONF by using the **show platform software yang-management process** command.

# Review Questions

1. True or false: NETCONF uses an RPC-based mechanism to facilitate client/server communication.
   - ❍ **A.** True
   - ❍ **B.** False

2. True or false: RESTCONF is based on JSON and SSH.
   - ❍ **A.** True
   - ❍ **B.** False

# Answers to Review Questions

1. **A** is correct. NETCONF uses a simple RPC-based mechanism to facilitate communication between a client (that is, a centralized management platform script or application) and a server (that is, a network device).
2. **B** is correct. RESTCONF is based on JSON and HTTP, whereas NETCONF-YANG uses an SSH model.

# Further Reading

- *CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide*
- *Network Programmability and Automation Fundamentals*

# What's Next?

If you want more practice on this chapter's exam objectives, remember that you can access all of the CramQuiz questions on the book's web page.

*This page intentionally left blank*

# Glossary

**802.1p** An IEEE specification that defines the use of the 3-bit Priority Code Point (PCP) field to provide different classes of service. The PCP field is contained within the TCI field, which is part of the 802.1Q header.

**802.1Q** An IEEE specification that defines two 2-byte fields—Tag Protocol Identifier (TPID) and Tag Control Information (TCI)—that are inserted within an Ethernet frame.

**802.1x** An IEEE standard for port-based network access control (PNAC) that provides an authentication mechanism for local area networks (LANs) and wireless LANs (WLANs).

## A

**access control list (ACL)** A mechanism that provides packet classification for quality of service (QoS), routing protocols, and basic firewall functionality.

**access layer** The network layer that gives endpoints and users direct access to the network.

**access port** A switch port that is configured for only one specific VLAN and generally connects end user devices.

**address family** A major classification of type of network protocol, such as IPv4, IPv6, or VPNv4.

**Address Resolution Protocol (ARP)** A protocol that resolves a MAC address to a specific IP address.

**administrative distance** A rating of trustworthiness for a route. Generally it is associated with the routing process that installs the route into the RIB.

**amplitude** The height from the top peak to the bottom peak of a signal's waveform; also known as the peak-to-peak amplitude.

**anchor controller** The original controller a client was associated with before a Layer 3 intercontroller roam. An anchor controller can also be used for tunneling clients on a guest WLAN or with a static anchor. Traffic is tunneled from the client's current controller (the foreign controller) back to the anchor.

**application programming interface (API)** A set of functions and procedures used for configuring or monitoring computer systems, network devices, or applications that involves programmatically interacting through software. Can be used for connecting to individual devices or multiple devices simultaneously.

**area border router (ABR)** A router that connects an OSPF area to Area 0 (that is, the backbone area).

**AS_Path** A BGP attribute used to track the autonomous systems a network has been advertised through as a loop-prevention mechanism.

**AS path access control list (ACL)** An ACL based on regex for identifying BGP routes based on the AS path and used for direct filtering or conditional matching in a route map.

**atomic aggregate** A BGP path attribute indicating that a prefix has been summarized, and not all of the path information from component routes was included in the aggregate.

**authentication, authorization, and accounting (AAA)** An architectural framework that enables secure network access control for users and devices.

**authentication server (AS)** An 802.1x entity that authenticates users or clients based on their credentials, as matched against a user database. In a wireless network, a RADIUS server is an AS.

**authenticator** An 802.1x entity that exists as a network device that provides access to the network. In a wireless network, a WLC acts as an authenticator.

**autonomous AP** A wireless AP operating in a standalone mode, such that it can provide a fully functional BSS and connect to the DS.

**autonomous system (AS)** A set of routers running the same routing protocol under a single realm of control and authority.

# B

**backbone area** The OSPF Area 0, which connects to all other OSPF areas. The backbone area is the only area that should provide connectivity between all other OSPF areas.

**backup designated router (BDR)** A backup pseudonode that maintains the network segment's state to replace the DR in the event of its failure.

**band** A contiguous range of frequencies.

**bandwidth** The range of frequencies used by a single channel or a single RF signal.

**beamwidth** A measure of the angle of a radiation pattern in both the E and the H planes, where the signal strength is 3 dB below the maximum value.

**BGP community** A well-known BGP attribute that allows for identification of routes for later actions such as identification of source or route filtering/modification.

**BGP multihoming** A method of providing redundancy and optimal routing that involves adding multiple links to external autonomous systems.

**BPDU filter** An STP feature that filters BPDUs from being advertised/received across the configured port.

**BPDU guard** An STP feature that places a port into an ErrDisabled state if a BPDU is received on a portfast-enabled port.

**bridge protocol data unit (BPDU)** A network packet that is used to identify a hierarchy and notify of changes in the topology.

**broadcast domain** A portion of a network where a single broadcast can be advertised or received.

**building block** A distinct place in the network (PIN) such as the campus end-user/endpoint block, the WAN edge block, the Internet edge block, or the network services block. The components of each building block are the access layer, the distribution layer, and/or the core (backbone) layer. Also known as a network block or a place in the network (PIN).

# C

**CAPWAP** A standards-based tunneling protocol that defines communication between a lightweight AP and a wireless LAN controller.

**carrier signal** The basic, steady RF signal that is used to carry other useful information.

**centralized WLC deployment** See unified WLC deployment.

**channel** An arbitrary index that points to a specific frequency within a band.

**Cisco Advanced Malware Protection (AMP)** A Cisco malware analysis and protection solution that goes beyond point-in-time detection and provides comprehensive protection for organizations across the full attack continuum: before, during, and after an attack.

**Cisco AnyConnect Secure Mobility Client** A VPN client that is an 802.1x supplicant that can perform posture validations and that provides web security, network visibility into endpoint flows within

Stealthwatch, and roaming protection with Cisco Umbrella.

**Cisco Email Security Appliance (ESA)** A Cisco solution that enables users to communicate securely via email and helps organizations combat email security threats with a multilayered approach across the attack continuum.

**Cisco Express Forwarding (CEF)** A method of forwarding packets in hardware through the use of the FIB and adjacency tables. CEF is much faster than process switching.

**Cisco Identity Services Engine (ISE)** A Cisco security policy management platform that provides highly secure network access control to users and devices across wired, wireless, and VPN connections. It allows for visibility into what is happening in the network, such as who is connected (endpoints, users, and devices), which applications are installed and running on endpoints (for posture assessment), and much more.

**Cisco SAFE** A framework that helps design secure solutions for the campus, data center, cloud, WAN, branch, and edge.

**Cisco Stealthwatch** A Cisco collector and aggregator of network telemetry data (NetFlow data) that performs network security analysis and monitoring to automatically detect threats that manage to infiltrate a network as well as threats that originate within a network.

**Cisco Talos** The Cisco threat intelligence organization.

**Cisco Threat Grid** A malware sandbox solution.

**Cisco TrustSec** A next-generation access control enforcement solution developed by Cisco that performs network enforcement by using Security Group Tags (SGTs) instead of IP addresses and ports. In SD-Access, Cisco TrustSec Security Group Tags are referred to as Scalable Group Tags.

**Cisco Umbrella** A Cisco solution that blocks requests to malicious Internet destinations (domains, IP addresses, URLs) using Domain Name System (DNS).

**Cisco Web Security Appliance (WSA)** An all-in-one web gateway that includes a wide variety of protections that can block hidden malware from both suspicious and legitimate websites.

**collision domain** A set of devices in a network that can transmit data packets that can collide with other packets sent by other devices (that is, devices that can detect traffic from other devices using CSMA/CD).

**command-line interface (CLI)** A text-based user interface for configuring network devices individually by inputting configuration commands.

**Common Spanning Tree (CST)** A single spanning-tree instance for the entire network, as defined in the 802.1D standard.

**configuration BPDU** The BPDU that is responsible for switches electing a root bridge and communicating the root path cost so that a hierarchy can be built.

**container** An isolated environment where containerized applications run. It contains the application along with the dependencies that the application needs to run. It is created by a container engine running a container image.

**container image** A file created by a container engine that includes application code along with its dependencies. Container images become containers when they are run by a container engine.

**content addressable memory (CAM)** A high-performance table used to correlate MAC addresses to switch interfaces that they are attached to.

**control plane policing (CoPP)** A policy applied to the control plane of a router to protect the CPU from high rates of traffic that could impact router stability.

**cookbook** A Chef container that holds recipes.

**core layer** The network layer, also known as the backbone, that provides high-speed connectivity between distribution layers in large environments.

# D

**Datagram Transport Layer Security (DTLS)** A communications protocol designed to provide authentication, data integrity, and confidentiality for communications between two applications, over a datagram transport protocol such as User Datagram Protocol (UDP). DTLS is based on TLS, and it includes enhancements such as sequence numbers and retransmission capability to compensate for the unreliable nature of UDP. DTLS is defined in IETF RFC 4347.

**dBd** dB-dipole, the gain of an antenna, measured in dB, as compared to a simple dipole antenna.

**dBi** dB-isotropic, the gain of an antenna, measured in dB, as compared to an isotropic reference antenna.

**dBm** dB-milliwatt, the power level of a signal measured in dB, as compared to a reference signal power of 1 milliwatt.

**dead interval** The amount of time required for a hello packet to be received for the neighbor to be deemed healthy. Upon receipt, the value resets and decrements toward zero.

**decibel (dB)** A logarithmic function that compares one absolute measurement to another.

**demodulation** The receiver's process of interpreting changes in the carrier signal to recover the original information being sent.

**designated port (DP)** A network port that receives and forwards BPDUs to other downstream switches.

**designated router (DR) (Context of OSPF)** A pseudonode to manage the adjacency state with other routers on the broadcast network segment.

**designated router (DR) (Context of PIM)** A PIM-SM router that is elected in a LAN segment when multiple PIM-SM routers exist to prevent the sending of duplicate multicast traffic into the LAN or the RP.

**DevNet** A single place to go to enhance or increase skills with APIs, coding, Python, and even controller concepts.

**Differentiated Services (DiffServ)** A field that uses the same 8 bits of the IP header that were previously used for the ToS and IPv6 Traffic Class fields. This allows it to be backward compatible with IP Precedence. The DiffServ field is composed of a 6-bit Differentiated Services Code Point (DSCP) field that allows for classification of up to 64 values (0 to 63) and a 2-bit Explicit Congestion Notification (ECN) field.

**Differentiated Services Code Point (DSCP)** A 6-bit field within the DiffServ field that allows for classification of up to 64 values (0 to 63).

**dipole** An omnidirectional antenna composed of two wire segments.

**direct sequence spread spectrum (DSSS)** A wireless LAN method in which a transmitter uses a single fixed, wide channel to send data.

**directional antenna** A type of antenna that propagates an RF signal in a narrow range of directions.

**directly attached static route** A static route that defines only the outbound interface for the next-hop device.

**discontiguous network** An OSPF network where Area 0 is not contiguous and generally results in routes not being advertised pervasively through the OSPF routing domain.

**distance vector routing protocol** A routing protocol that selects the best path based on next hop and hop count.

**distribute list** A list used for filtering routes with an ACL for a specific BGP neighbor.

**distribution layer** The network layer that provides an aggregation point for the access layer and acts as a services and control boundary between the access layer and the core layer.

**downstream** Away from the source of a tree and toward the receivers.

**downstream interface** An interface that is used to forward multicast traffic down the tree, also known as the outgoing interface (OIF).

**dynamic rate shifting (DRS)** A mechanism used by an 802.11 device to change the modulation coding scheme (MCS) according to dynamic RF signal conditions.

**Dynamic Trunking Protocol (DTP)** A protocol that allows for the dynamic negotiation of trunk ports.

# E

**E plane** The "elevation" plane, which passes through an antenna that shows a side view of the radiation pattern.

**eBGP session** A BGP session maintained with BGP peers from a different autonomous system.

**effective isotropic radiated power (EIRP)** The resulting signal power level, measured in dBm, of the combination of a transmitter, cable, and an antenna, as measured at the antenna.

**egress tunnel router (ETR)** A router that de-encapsulates LISP-encapsulated IP packets coming from other sites and destined to EIDs within a LISP site.

**Embedded Event Manager (EEM)** An on-box automation tool that allows scripts to automatically execute, based on the output of an action or an event on a device.

**embedded WLC deployment** A wireless network design that places a WLC in the access layer, co-located with a LAN switch stack, near the APs it controls.

**endpoint** A device that connects to a network, such as a laptop, tablet, IP phone, personal computer (PC), or Internet of Things (IoT) device.

**endpoint identifier (EID)** The IP address of an endpoint within a LISP site.

**enhanced distance vector routing protocol** A routing protocol that selects the best path based on next hop, hop count, and other metrics, such as bandwidth and delay.

**equal-cost multipathing** The installation of multiple best paths from the same routing protocol with the same metric that allows for load-balancing of traffic across the paths.

**ERSPAN** Encapsulated Remote Switched Port Analyzer, a tool for capturing network traffic on a remote device and sending the traffic to the local system via Layer 3 (routing) toward a local port that would be attached to some sort of traffic analyzer.

**EtherChannel bundle** A logical interface that consists of physical member links to increase a link's bandwidth while preventing forwarding loops.

**Extensible Authentication Protocol (EAP)** A standardized authentication framework defined by RFC 4187 that provides encapsulated transport for authentication parameters.

**Extensible Markup Language (XML)** A human-readable data format that is commonly used with web services.

# F

**feasibility condition** A condition under which, for a route to be considered a backup route, the reported distance received for that route must be less than the feasible distance calculated locally. This logic guarantees a loop-free path.

**feasible distance** The metric value for the lowest-metric path to reach a destination.

**feasible successor** A route that satisfies the feasibility condition and is maintained as a backup route.

**first-hop redundancy protocol** A protocol that creates a virtual IP address on a router or a multi-layer device to ensure continuous access to a gateway when there are redundant devices.

**first-hop router (FHR)** A router that is directly attached to the source, also known as the root router. It is responsible for sending register messages to the RP.

**floating static route** A static route with an elevated AD so that it is used only as a backup in the event that a routing protocol fails or a lower-AD static route is removed from the RIB.

**foreign controller** The current controller that a client is associated with after a Layer 3 intercontroller roam. Traffic is tunneled from the foreign controller back to an anchor controller so that the client retains connectivity to its original VLAN and subnet.

**forward delay** The amount of time that a port stays in a listening and learning state.

**Forwarding Information Base (FIB)** The hardware programming of a forwarding table. The FIB uses the RIB for programming.

**frequency** The number of times a signal makes one complete up and down cycle in 1 second.

**fully specified static route** A static route that specifies the next-hop IP address and the outbound interface.

# G

**gain** A measure of how effectively an antenna can focus RF energy in a certain direction.

**GitHub** An efficient and commonly adopted way of using version control for code and sharing code repositories.

**grain** In SaltStack, code that runs on nodes to gather system information and report back to the master.

# H

**H plane** The "azimuth" plane, which passes through an antenna that shows a top-down view of the radiation pattern.

**hello interval** The frequency at which hello packets are advertised out an interface.

**hello packets** Packets that are sent out at periodic intervals to detect neighbors for establishing adjacency and ensuring that neighbors are still available.

**hello time** The time interval for which a BPDU is advertised out of a port.

**hello timer** The amount of time between the advertisement of hello packets and when they are sent out an interface.

**hertz (Hz)** A unit of frequency equaling one cycle per second.

**host pool** The IP subnet, SVI, and VRF information assigned to a group of hosts that share the same policies.

**hypervisor** Virtualization software that creates VMs and performs the hardware abstraction that allows multiple VMs to run concurrently.

# I

**iBGP session** A BGP session maintained with BGP peers from the same autonomous system.

**IGMP snooping** A mechanism to prevent multicast flooding on a Layer 2 switch.

**in phase** The condition when the cycles of two identical signals are in sync with each other.

**incoming interface (IIF)** The only type of interface that can accept multicast traffic coming from the source. It is the same as the RPF interface.

**ingress tunnel router (ITR)** A router that LISP-encapsulates IP packets coming from EIDs that are destined outside the LISP site.

**inside global** The public IP address that represents one or more inside local IP addresses to the outside.

**inside local** The actual private IP address assigned to a device on the inside network(s).

**integrated antenna** A very small omnidirectional antenna that is set inside a device's outer case.

**interarea route** An OSPF route learned from an ABR from another area. These routes are built based on type 3 LSAs.

**intercontroller roaming** Client roaming that occurs between two APs that are joined to two different controllers.

**interface priority** The reference value for an interface to determine preference for being elected as the designated router.

**internal spanning tree (IST)** The first MSTI in the MST protocol. The IST is responsible for building a CST across all VLANs, regardless of their VLAN membership. The IST contains advertisements for other MSTIs in its BPDUs.

**Internet Group Management Protocol (IGMP)** The protocol used by receivers to join multicast groups and start receiving traffic from those groups.

**Internet Key Exchange (IKE)** A protocol that performs authentication between two endpoints to establish security associations (SAs), also known as IKE tunnels. IKE is the implementation of ISAKMP using the Oakley and Skeme key exchange techniques.

**Internet Protocol Security (IPsec)** A framework of open standards for creating highly secure VPNs using various protocols and technologies for secure communication across unsecure networks such as the Internet.

**Internet Security Association Key Management Protocol (ISAKMP)** A framework for authentication and key exchange between two peers to establish, modify, and tear down SAs that is designed to support many different kinds of key exchanges.

ISAKMP uses UDP port 500 to communicate between peers.

**intra-area route** An OSPF route learned from a router within the same area. These routes are built based on type 1 and type 2 LSAs.

**intracontroller roaming** Client roaming that occurs between two APs joined to the same controller.

**IP SLA** An on-box diagnostic tool that executes probes to monitor network devices and application performance.

**isotropic antenna** An ideal, theoretical antenna that radiates RF equally in every direction.

# J

**JavaScript Object Notation (JSON)** Notation used to store data in key/value pairs that is said to be easier to work with and read than XML.

# K

**K values** Values that EIGRP uses to calculate the best path.

# L

**LACP interface priority** An attribute assigned to a switch port on an LACP primary switch to identify which member links are used when there is a maximum link.

**LACP system priority** An attribute in an LACP packet that provides priority to one switch over another to control which links are used when there is a maximum link.

**last-hop router (LHR)** A router that is directly attached to the receivers, also known as a leaf router. It is responsible for sending PIM joins upstream toward the RP or to the source after an SPT switchover.

**Layer 2 forwarding** The forwarding of packets based on the packets' destination Layer 2 addresses, such as MAC addresses.

**Layer 2 roam** An intercontroller roam where the WLANs of the two controllers are configured for the same Layer 2 VLAN ID; also known as a local-to-local roam.

**Layer 3 forwarding** The forwarding of packets based on the packets' destination IP addresses.

**Layer 3 roam** An intercontroller roam where the WLANs of the two controllers are configured for different VLAN IDs; also known as a local-to-foreign roam. To support the roaming client, a tunnel is built between the controllers so that client data can pass between the client's current controller and its original controller.

**lightweight AP** A wireless AP that performs real-time 802.11 functions to interface with wireless clients, while relying on a wireless LAN controller to handle all management functions.

**link budget** The cumulative sum of gains and losses measured in dB over the complete RF signal path; a transmitter's power level must overcome the link budget so that the signal can reach a receiver effectively.

**link-state routing protocol** A routing protocol that contains a complete view of the topology, where every router can calculate the best path based on its copy of the topology.

**LISP router** A router that performs the functions of any or all of the following: ITR, ETR, PITR, and/or PETR.

**LISP site** A site where LISP routers and EIDs reside.

**load-balancing hash** An algorithm for balancing network traffic across member links.

**Loc-RIB table** The main BGP table that contains all the active BGP prefixes and path attributes that is used to select the best path and install routes into the RIB.

**local bridge identifier** A combination of the advertising switch's bridge system MAC, the system ID extension, and the system priority of the local bridge.

**local mode** The default mode of a Cisco lightweight AP that offers one or more functioning BSSs on a specific channel.

**Location/ID Separation Protocol (LISP)** A routing architecture and data and control plane protocol that was created to address routing scalability problems on large networks.

# M

**MAC address table** A table on a switch that identifies the switch port and VLAN with which a MAC address is associated for Layer 2 forwarding.

**MAC Authentication Bypass (MAB)** A network access control technique that enables port-based access control using the MAC address of an endpoint and is typically used as a fallback mechanism to 802.1x.

**MACsec** An IEEE 802.1AE standards-based Layer 2 link encryption technology used by TrustSec to encrypt Secure Group Tag (SGT) frames on Layer 2 links between switches and between switches and endpoints.

**manifest** In Puppet, the code to be executed that is contained within modules.

**map resolver (MR)** A network device (typically a router) that receives LISP-encapsulated map requests from an ITR and finds the appropriate ETR to answer those requests by consulting the map server. If requested by the ETR, the MS can reply on behalf of the ETR.

**map server (MS)** A network device (typically a router) that learns EID-to-prefix mapping entries from an ETR and stores them in a local EID-to-RLOC mapping database.

**map server/map resolver (MS/MR)** A device that performs MS and MR functions. The MS function learns EID-to-prefix mapping entries from an ETR and stores

them in a local EID-to-RLOC mapping database. The MR function receives LISP-encapsulated map requests from an ITR and finds the appropriate ETR to answer those requests by consulting the mapping server. If requested by the ETR, the MS can reply on behalf of the ETR.

**max age** The timer that controls the maximum length of time that passes before a bridge port saves its BPDU information.

**maximal-ratio combining (MRC)** An 802.11n technique that combines multiple copies of a signal, received over multiple antennas, to reconstruct the original signal.

**member links** The physical links used to build a logical EtherChannel bundle.

**mobility domain** A logical grouping of all mobility groups within an enterprise.

**Mobility Express WLC deployment** A wireless network design that places a WLC co-located with a lightweight AP.

**mobility group** A logical grouping of one or more MCs between which efficient roaming is expected.

**modulation** The transmitter's process of altering the carrier signal according to some other information source.

**module** A Puppet container that holds manifests.

**MST instance (MSTI)** A single spanning-tree instance for a specified set of VLANs in the MST protocol.

**MST region** A collection of MSTIs that operate in the same MST domain.

**MST region boundary** Any switch port that connects to another switch in a different MST region or that connects to a traditional 802.1D or 802.1W STP instance.

**Multicast Forwarding Information Base (MFIB)** A forwarding table that derives information from the MRIB to program multicast forwarding information in hardware for faster forwarding.

**Multicast Routing Information Base (MRIB)** A topology table that is also known as the multicast route table (mroute), which derives from the unicast routing table and PIM.

**multicast state** The traffic forwarding state that is used by a router to forward multicast traffic. The multicast state is composed of the entries found in the mroute table (S, G, IIF, OIF, and so on).

# N

**narrowband** RF signals that use a very narrow range of frequencies.

**native VLAN** A VLAN that correlates to any untagged network traffic on a trunk port.

**NETCONF** A protocol defined by the IETF for installing, manipulating, and deleting the configuration of network devices.

**NetFlow** A Cisco network protocol for exporting flow information generated from network devices in order to analyze traffic statistics.

**Network Address Translation (NAT)** The systematic modification of source and/or destination IP headers on a packet from one IP address to another.

**network block** See building block.

**Network Configuration Protocol (NETCONF)/YANG** An IETF standard protocol that uses the YANG data models to communicate with the various devices on the network. NETCONF runs over SSH, TLS, or Simple Object Access Protocol (SOAP).

**network function (NF)** The function performed by a physical appliance, such as a firewall function or a router function.

**network functions virtualization (NFV)** An architectural framework created by the European Telecommunications Standards Institute (ETSI) that defines standards to decouple network functions from proprietary hardware-based appliances and have them run in software on standard x86 servers.

**network LSA** A type 2 LSA that advertises the routers connected to the DR pseudonode. Type 2 LSAs remain within the OSPF area of origination.

**next-generation firewall (NGFW)** A firewall with legacy firewall capabilities such as stateful inspection as well as integrated intrusion prevention, application-level inspection, and techniques to address evolving security threats, such as advanced malware and application-layer attacks.

**NFV infrastructure (NFVI)** All the hardware and software components that comprise the platform environment in which virtual network functions (VNFs) are deployed.

**noise floor** The average power level of noise measured at a specific frequency.

**nonce** A random or pseudo-random number issued in an authentication protocol that can be used just once to prevent replay attacks.

**NTP client** A device that queries a time server by using Network Time Protocol so that it can synchronize its time to the server.

**NTP peer** A device that queries another peer device using Network Time Protocol so that the two devices can synchronize and adjust their time to each other.

**NTP server** A device that provides time to clients that query it with Network Time Protocol.

# O

**omnidirectional antenna** A type of antenna that propagates an RF signal in a broad range of directions in order to cover a large area.

**Open Authentication** An 802.11 authentication method that requires clients to associate with an AP without providing any credentials.

**optional non-transitive** A BGP path attribute that might be recognized by a BGP implementation that is not advertised between autonomous systems.

**optional transitive** A BGP path attribute that might be recognized by a BGP implementation that is advertised between autonomous systems.

**Orthogonal Frequency Division Multiplexing (OFDM)** A data transmission method that sends data bits in parallel over multiple frequencies within a single 20 MHz wide channel. Each frequency represents a single subcarrier.

**out of phase** The condition when the cycles of one signal are shifted in time in relation to another signal.

**outgoing interface (OIF)** An interface that is used to forward multicast traffic down the tree, also known as the downstream interface.

**outgoing interface list (OIL)** A group of OIFs that are forwarding multicast traffic to the same group.

**outside global** The public IP address assigned to a host on the outside network by the owner of the host. This IP address must be reachable by the outside network.

**outside local** The IP address of an outside host as it appears to the inside network. The IP address does not have to be reachable by the outside but is considered private and must be reachable by the inside network.

**overlay network** A logical or virtual network built over a physical transport network referred to as an underlay network.

# P

**parabolic dish antenna** A highly directional antenna that uses a passive dish shaped like a parabola to focus an RF signal into a tight beam.

**passive interface** An interface that has been enabled with a routing protocol to advertise its associated interfaces into its RIB but that does not establish neighborship with other routers associated to that interface.

**patch antenna** A directional antenna that has a planar surface and is usually mounted on a wall or column.

**Path Trace** A visual troubleshooting tool in Cisco DNA Center Assurance that is used to trace a route and display the path throughout the network between wired or wireless hosts.

**path vector routing protocol** A routing protocol that selects the best path based on path attributes.

**per-hop behavior (PHB)** The QoS action applied to a packet (expediting, delaying, or dropping) on a hop-by-hop basis, based on its DSCP value.

**personal mode** Pre-Shared Key authentication as applied to WPA, WPA2, or WPA3.

**phase** A measure of shift in time relative to the start of a cycle; ranges between 0 and 360 degrees.

**pillar** A SaltStack value store that stores information that a minion can access from the master.

**place in the network (PIN)** See building block.

**play** In Ansible, the code to be executed that is contained within playbooks.

**playbook** An Ansible container that holds plays.

**polar plot** A round graph that is divided into 360 degrees around an antenna and into concentric circles that represent decreasing dB values. The antenna is always placed at the center of the plot.

**polarization** The orientation (horizontal, vertical, circular, and so on) of a propagating wave with respect to the ground.

**pooled NAT** A dynamic one-to-one mapping of a local IP address to a global IP address. The global IP address is temporarily assigned to a local IP address. After a certain amount of idle NAT time, the global IP address is returned to the pool.

**Port Address Translation (PAT)** A dynamic many-to-one mapping of a global IP address to many local IP addresses. The NAT device keeps track of the global IP address-to-local IP address mappings using multiple port numbers.

**prefix length** The number of leading binary bits in the subnet mask that are in the on position.

**prefix list** A method of selecting routes based on binary patterns, specifically the high-order bit pattern, high-order bit count, and an optional prefix length parameter.

**privilege level** A Cisco IOS CLI designation of what commands are available to a user.

**process switching** The process of forwarding traffic by software and processing by the general CPU. It is typically slower than hardware switching.

**Protocol Independent Multicast (PIM)** A multicast routing protocol that routes multicast traffic between network segments. PIM can use any of the unicast routing protocols to identify the path between the source and receivers.

**proxy ETR (PETR)** An ETR but for LISP sites that sends traffic to destinations at non-LISP sites.

**proxy ITR (PITR)** An ITR but for a non-LISP site that sends traffic to EID destinations at LISP sites.

**proxy xTR (PxTR)** A router that performs proxy ITR (PITR) and proxy ETR (PETR) functions.

**PVST simulation check** The process of ensuring that the MST region is the STP root bridge for all the VLANs or none of the VLANs. If the MST region is a partial STP root bridge, the port is shut down.

**Python** A commonly used programming language that is easy to interpret and use. It is often used to manage network devices and for software scripting.

# Q

**quadrature amplitude modulation (QAM)** A modulation method that combines QPSK phase shifting with multiple amplitude levels to produce a greater number of unique changes to the carrier signal. The number preceding the QAM name designates how many carrier signal changes are possible.

# R

**radiation pattern** A plot that shows the relative signal strength in dBm at every angle around an antenna.

**radio frequency (RF)** The portion of the frequency spectrum between 3 kHz and 300 GHz.

**RADIUS server** An authentication server used with 802.1x to authenticate wireless clients.

**received signal strength (RSS)** The signal strength level in dBm that an AP receives from a wireless device.

**received signal strength indicator (RSSI)** The relative measure of signal strength (0 to 255), as seen by the receiver.

**recipe** In Chef, the code to be executed that is contained within cookbooks.

**recursive static route** A static route that specifies the next-hop IP address and requires the router to recursively locate the outbound interface for the next-hop device.

**regular expressions (regex)** Search patterns that use special key characters for parsing and matching.

**Remote Authentication Dial-In User Service (RADIUS)** AAA protocol that is primarily used to enable network access control (secure access to network resources).

**rendezvous point (RP)** A single common root placed at a chosen point of a shared distribution tree. In other words, it is the root of a shared distribution tree known as a rendezvous point tree (RPT).

**rendezvous point tree (RPT)** Also known as a shared tree, a multicast distribution tree where the root of the shared tree is not the source but a router designated as the rendezvous point (RP).

**reported distance** The distance reported by a router to reach a prefix. The reported distance value is the feasible distance for the advertising router.

**RESTCONF** An IETF draft that describes how to map a YANG specification to a RESTful interface.

**Reverse Path Forwarding (RPF) interface** The interface with the lowest-cost path (based on administrative distance [AD] and metric) to the IP address of the source (SPT) or the RP.

**RF fingerprinting** A method used to accurately determine wireless device location by applying a calibration model to the location algorithm so that the RSS values measured also reflect the actual environment.

**root bridge** The topmost switch in an STP topology. The root bridge is responsible for controlling STP timers, creating configuration

BPDUs, and processing topology change BPDUs. All ports on a root bridge are designated ports that are in a forwarding state.

**root bridge identifier** A combination of the root bridge system MAC address, system ID extension, and system priority of the root bridge.

**root guard** An STP feature that places a port into an ErrDisabled state if a superior BPDU is received on the configured port.

**root path cost** The cost for a specific path toward the root switch.

**root port** The most preferred switch port that connects a switch to the root bridge. Often this is the switch port with the lowest root path cost.

**route map** A feature used in BGP (and other IGP components) that allows for filtering or modification of routes using a variety of conditional matching.

**router ID (RID)** A 32-bit number that uniquely identifies the router in a routing domain.

**router LSA** A type 1 LSA that is a fundamental building block representing an OSPF-enabled interface. Type 1 LSAs remain within the OSPF area of origination.

**Routing Information Base (RIB)** The software database of all the routes, next-hop IP addresses, and attached interfaces. Also known as a routing table.

**routing locator (RLOC)** An IPv4 or IPv6 address of an ETR that is Internet facing or network core facing.

**RPF neighbor** The PIM neighbor on the RPF interface.

**RSPAN** Remote Switched Port Analyzer, a tool for capturing network traffic on a remote switch and sending a copy of the network traffic to the local switch via Layer 2 (switching) toward a local port that would be attached to some sort of traffic analyzer.

# S

**Scalable Group Tag (SGT)** A technology that is used to perform ingress tagging and egress filtering to enforce access control policy. The SGT tag assignment is delivered to the authenticator as an authorization option. After the SGT tag is assigned, an access enforcement policy based on the SGT tag can be applied at any egress point of the TrustSec network. In SD-Access, Cisco TrustSec Security Group Tags are referred to as Scalable Group Tags.

**Secure Shell (SSH)** A secure network communication protocol that provides secure encryption and strong authentication.

**Security Group Access Control List (SGACL)** A technology that provides filtering based on source and destination SGT tags.

**segment** An overlay network.

**segmentation** A process that enables a single network infrastructure to support multiple Layer 2 or Layer 3 overlay networks.

**sensitivity level** The RSSI threshold (in dBm) that divides unintelligible RF signals from useful ones.

**service chaining** Chaining VNFs together to provide an NFV service or solution.

**shortest path tree (SPT)** A router's view of the topology to reach all destinations in the topology, where the router is the top of the tree, and all of the destinations are the branches of the tree. In the context of multicast, the SPT provides a multicast distribution tree where the source is the root of the tree and branches form a distribution tree through the network all the way down to the receivers. When this tree is built, it uses the shortest path through the network from the source to the leaves of the tree.

**signal-to-noise ratio (SNR)** A measure of received signal quality, calculated as the difference between the signal's RSSI and the noise floor. A higher SNR is preferred.

**Simple Network Management Protocol (SNMP)** A protocol that can send alerts when something fails on a device as well as when certain events happen on a device (for example, power supply failure).

**SPAN** Switched Port Analyzer, a tool for capturing local network traffic on a switch and sending a copy of the network traffic to a local port that would be attached to some sort of traffic analyzer.

**spatial multiplexing** Distributing streams of data across multiple radio chains with spatial diversity.

**spatial stream** An independent stream of data that is sent over a radio chain through free space. One spatial stream is separate from others due to the unique path it travels through space.

**split-MAC architecture** A wireless AP strategy based on the idea that normal AP functions are split or divided between a wireless LAN controller and lightweight APs.

**spread spectrum** RF signals that spread the information being sent over a wide range of frequencies.

**static NAT** A static one-to-one mapping of a local IP address to a global IP address.

**static null route** A static route that specifies the virtual null interface as the next hop as a method of isolating traffic or preventing routing loops.

**STP loop guard** An STP feature that prevents a configured alternative or root port from becoming a designated port toward a downstream switch.

**STP portfast** An STP feature that places a switch port directly into a forwarding state and disables TCN generation for a change in link state.

**stratum** A level that makes it possible to identify the accuracy of the time clock source, where the lower the stratum number, the more accurate the time is considered.

**successor** The first next-hop router for the successor route.

**successor route** The route with the lowest path metric to reach a destination.

**summarization** A method of reducing a routing table by advertising a less specific network prefix in lieu of multiple more specific network prefixes.

**summary LSA** A type 3 LSA that contains the routes learned from another area. Type 3 LSAs are generated on ABRs.

**supplicant** An 802.1x entity that exists as software on a client device and serves to request network access.

**syslog** Logging of messages that can be sent to a collector server or displayed on the console or stored in the logging buffer on the local device.

**system ID extension** A 12-bit value that indicates the VLAN that the BPDU correlates to.

**system priority** A 4-bit value that indicates the preference for a switch to be root bridge.

# T

**Tcl** A scripting language that can be run on Cisco IOS devices to automate tasks such as ping scripts.

**Telnet** An insecure network communication protocol that communicates using plaintext and is not recommended for use in production environments.

**Terminal Access Controller Access-Control System Plus (TACACS+)** AAA protocol that is primarily used to enable device access control (secure access to network devices).

**ternary content addressable memory (TCAM)** A high-performance table or tables that can evaluate packet forwarding decisions based on policies or access lists.

**topology change notification (TCN)** A BPDU that is advertised toward the root bridge to notify the root of a topology change on a downstream switch.

**topology table** A table used by EIGRP that maintains all network prefixes, advertising EIGRP neighbors for prefixes and path metrics for calculating the best path.

**transit routing** The act of allowing traffic to flow from one external autonomous system through your autonomous system to reach a different external autonomous system.

**transmit beamforming (T×BF)** A method of transmitting a signal over multiple antennas, each having the signal phase carefully crafted, so that the multiple copies are all in phase at a targeted receiver.

**trunk port** A switch port that is configured for multiple VLANs and generally connects a switch to other switches or to other network devices, such as firewalls or routers.

**tunnel router (xTR)** A router that performs ingress tunnel router (ITR) and egress tunnel router (ETR) functions (which is most routers).

**Type of Service (TOS)** An 8-bit field where only the first 3 bits, referred to as IP Precedence (IPP), are used for marking, and the rest of the bits are unused. IPP values range from 0 to 7 and allow the

traffic to be partitioned into up to six usable classes of service; IPP 6 and 7 are reserved for internal network use.

# U

**underlay network** The traditional physical networking infrastructure that uses an IGP or a BGP.

**unequal-cost load balancing** The installation of multiple paths that include backup paths from the same routing protocol. Load balancing across the interface uses a traffic load in a ratio to the interface's route metrics.

**Unidirectional Link Detection (UDLD)** A protocol that provides bidirectional monitoring of fiber-optic cables.

**unified WLC deployment** A wireless network design that places a WLC centrally within a network topology.

**upstream** Toward the source of a tree, which could be the actual source with a source-based tree or the RP with a shared tree. A PIM join travels upstream toward the source.

**upstream interface** The interface toward the source of the tree. Also known as the RPF interface or the incoming interface (IIF).

# V

**variance value** The feasible distance (FD) for a route multiplied by the EIGRP variance multiplier. Any feasible successor's FD with a metric below the EIGRP variance value is installed into the RIB.

**virtual local area network (VLAN)** A logical segmentation of switch ports based on the broadcast domain.

**virtual machine (VM)** A software emulation of a physical server with an operating system.

**virtual network (VN)** Virtualization at the device level, using virtual routing and forwarding (VRF) instances to create multiple Layer 3 routing tables.

**virtual network function (VNF)** The virtual version of an NF, typically run on a hypervisor as a VM (for example, a virtual firewall such as the ASAv or a virtual router such as the ISRv).

**virtual private network (VPN)** An overlay network that allows private networks to communicate with each other across an untrusted underlay network such as the Internet.

**virtual switch (vSwitch)** A software-based Layer 2 switch that operates like a physical Ethernet switch and enables VMs to communicate with each other within a virtualized server and with external physical networks using physical network interface cards (pNICs).

**virtual tunnel endpoint (VTEP)** An entity that originates or terminates a VXLAN tunnel. It maps Layer 2 and Layer 3 packets to the VNI to be used in the overlay network.

**VLAN Trunking Protocol (VTP)** A protocol that enables the provisioning of VLANs on switches.

**VXLAN** An overlay data plane encapsulation scheme that was developed to address the various issues seen in traditional Layer 2 networks. It does this by extending Layer 2 and Layer 3 overlay networks over a Layer 3 underlay network, using MAC-in-IP/UDP tunneling. Each overlay is termed a VXLAN segment.

**VXLAN Group Policy Option (GPO)** An enhancement to the VXLAN header that adds new fields to the first 4 bytes of the VXLAN header in order to support and carry up to 64,000 SGT tags.

**VXLAN network identifier (VNI)** A 24-bit field in the VXLAN header that enables up to 16 million Layer 2 and/or Layer 3 VXLAN segments to coexist within the same infrastructure.

# W–X

**wavelength** The physical distance that a wave travels over one complete cycle.

**Web Authentication (WebAuth)** A network access control technique that enables access control by presenting a guest web portal requesting a username and password. It is typically used as a fallback mechanism to 802.1x and MAB.

**well-known discretionary** A BGP path attribute recognized by all BGP implementations that may or may not be advertised to other peers.

**well-known mandatory** A BGP path attribute recognized by all BGP implementations that must be advertised to other peers.

**wide metrics** A new method of advertising and identifying interface speeds and delay to account for higher-bandwidth interfaces (20 Gbps and higher).

**Wi-Fi Protected Access (WPA)** A Wi-Fi Alliance standard that requires pre-shared key or 802.1x authentication, TKIP, and dynamic encryption key management; based on portions of 802.11i before its ratification.

**wireless LAN controller (WLC)** A device that controls and manages multiple lightweight APs.

**WPA Version 2 (WPA2)** A Wi-Fi Alliance standard that requires Pre-Shared Key or 802.1x authentication, TKIP or CCMP, and dynamic encryption key management; based on the complete 802.11i standard after its ratification.

**WPA Version 3 (WPA3)** The third version of a Wi-Fi Alliance standard, introduced in 2018, that requires Pre-Shared Key or 802.1x authentication, GCMP, SAE, and forward secrecy.

# Y

**Yagi antenna** A directional antenna made up of several parallel wire segments that tend to amplify an RF signal to each other.

**YANG Model** A model that represents anything that can be configured or monitored, as well as all administrative actions that can be taken on a device.

# Z

**Zone Based Firewall (ZBFW)** An IOS integrated stateful firewall.

# Index

## Symbols

# A

## B

## C

# D

## E

# F

## G

## H

## I

## J



## K



## L

## M

# N

## O

## P

## Q

## R

# S

# T

## U

## V

## W

## X



## Y



## Z

*This page intentionally left blank*

Photo by Olena Yakobchuk/Shutterstock

## Register Your Product at pearsonITcertification.com/register

Access additional benefits and **save 35%** on your next purchase

- Automatically receive a coupon for 35% off your next purchase, valid for 30 days. Look for your code in your Pearson IT Certification cart or the Manage Codes section of your account page.
- Download available product updates.
- Access bonus material if available.*
- Check the box to hear from us and receive exclusive offers on new editions and related products.

**Registration benefits vary by product. Benefits will be listed on your account page under Registered Products.*

---

## Learning Solutions for Self-Paced Study, Enterprise, and the Classroom

Pearson IT Certification delivers training materials that address the learning, preparation, and practice needs of a new generation of certification candidates, including the official publishing programs of Adobe Press, Cisco Press, and Microsoft Press. At pearsonITcertification.com, you can:

- Shop our books, eBooks, practice tests, software, and video courses
- Sign up to receive special offers
- Access thousands of free chapters and video lessons

Visit **pearsonITcertification.com/community** to connect with Pearson IT Certification

Pearson

Addison-Wesley • Adobe Press • Cisco Press • Microsoft Press • Pearson IT Certification • Que • Sams • Peachpit Press

# CCNP® and CCIE® Enterprise Core

## ENCOR 350-401
## Companion Website

---

Access interactive study tools on this book's companion website, including practice test software, Glossary, and Cram Sheet.

To access the companion website, simply follow these steps:

1. Go to **www.pearsonitcertification.com/register**.
2. Enter the print book ISBN: **9780136891932**.
3. Answer the security question to validate your purchase.
4. Go to your account page.
5. Click on the Registered Products tab.
6. Under the book listing, click on the Access Bonus Content link.

If you have any issues accessing the companion website, you can contact our support team by going to **http://pearsonitp.echelp.org**.

Pearson

# Where are the companion content files?

## Register this digital version of CCNP and CCIE Enterprise Core ENCOR 350-401 Exam Cram to access important downloads.

Register this eBook to unlock the companion files. Follow these steps:

1. Go to **pearsonITcertification.com/account** and log in or create a new account.
2. Enter the ISBN: **9780136891932** (NOTE: Please enter the print book ISBN provided to register the eBook you purchased.)
3. Answer the challenge question as proof of purchase.
4. Click on the "Access Bonus Content" link in the Registered Products section of your account page, to be taken to the page where your downloadable content is available.

This eBook version of the print title does not contain the practice test software that accompanies the print book.

You May Also Like—Premium Edition eBook and Practice Test. To learn about the Premium Edition eBook and Practice Test series, visit **pearsonITcertification.com/practicetest**

---

The Professional and Personal Technology Brands of Pearson

Addison Wesley · Cisco Press · informIT · PEARSON IT Certification · que · SAMS