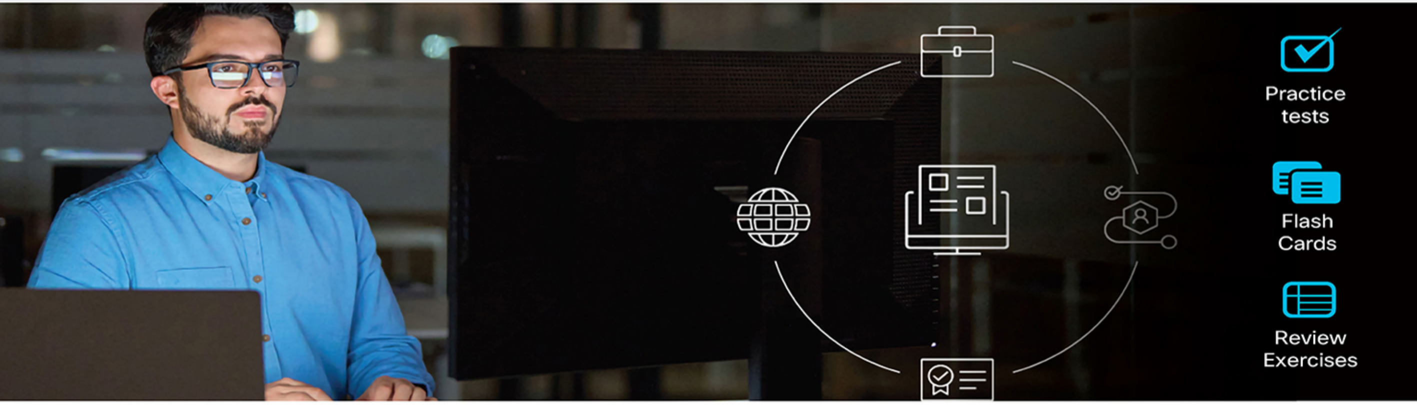
Official Cert Guide



Study Planner

CCNP

SPCOR 350-501

BRADLEY RIAPOLOV, CCIE® No. 18921
MOHAMMAD S. KHALIL, CCIE® No. 35484

ciscopress.com

CCNP SPCOR 350-501 Official Cert Guide

Companion Website and Pearson Test Prep Access Code

Access interactive study tools on this book's companion website, including practice test software, review exercises, Key Term flash card application, a study planner, and more!

To access the companion website, simply follow these steps:

1. Go to www.ciscopress.com/register by December 31, 2027.
2. Enter the **print book ISBN**: 9780135324806.
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.

When you register your book, your Pearson Test Prep practice test access code will automatically be populated with the book listing under the Registered Products tab. You will need this code to access the practice test that comes with this book. You can redeem the code at **PearsonTestPrep.com**. Simply choose Pearson IT Certification as your product group and log into the site with the same credentials you used to register your book. Click the **Activate New Product** button and enter the access code. More detailed instructions on how to redeem your access code for both the online and desktop versions can be found on the companion website.

If you have any issues accessing the companion website or obtaining your Pearson Test Prep practice test access code, you can contact our support team by going to **pearsonitp.echelp.org**.

This page intentionally left blank

CCNP SPCOR 350-501

Official Cert Guide

BRADLEY RIAPOLOV, CCIE No. 18921

MOHAMMAD KHALIL, CCDE No. 2023::57, CCIE No. 35484

Cisco Press

CCNP SPCOR 350-501 Official Cert Guide

Bradley Riapolov

Mohammad Khalil

Published by:
Cisco Press
Hoboken, New Jersey

$PrintCode

Library of Congress Control Number: 2024947468

ISBN-13: 978-0-13-532480-6

ISBN-10: 0-13-532480-7

Warning and Disclaimer

This book is designed to provide information about the Implementing and Operating Cisco Service Provider Network Core Technologies (SPCOR 350-501) exam. Every effort has been made to make this book as complete and as accurate as possible, but no warranty or fitness is implied.

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

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

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

Trademark Acknowledgments

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

Feedback Information

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

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

We greatly appreciate your assistance.

GM K12, Early Career and Professional Learning: Soo Kang

Alliances Manager, Cisco Press: Caroline Antonio

Director, ITP Product Management: Brett Bartow

Executive Editor: James Manly

Managing Editor: Sandra Schroeder

Development Editor: Christopher A. Cleveland

Senior Project Editor: Tonya Simpson

Copy Editor: CAH Editorial

Technical Editor: Brad Edgeworth

Editorial Assistant: Cindy Teeters

Cover Designer: Chuti Prasertsith

Composition: codeMantra

Indexer: Timothy Wright

Proofreader: Jennifer Hinchliffe

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV
Amsterdam, The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

About the Authors

Bradley Riapolov is a seasoned technical solutions architect at Cisco Systems, currently serving in the MassScale Infrastructure Group. Since joining Cisco in 2008, he has been at the forefront of designing and deploying cutting-edge networking solutions, leveraging his extensive expertise to meet the complex demands of today's technology landscape.

Before his tenure at Cisco, Bradley gained invaluable experience working with various Fortune 500 companies, where he was instrumental in designing and operating small, medium, and large networks. His 20-year career is marked by a diverse background in successfully implementing technical campaigns across multiple industries, including transport, service provider, enterprise, industrial, and mobility networking.

In addition to his role at Cisco, Bradley is a recognized thought leader and educator. As a Cisco Press author and a distinguished speaker at Cisco Live, he has contributed to the body of knowledge in networking, sharing his insights and expertise through various publications. Moreover, his dedication to education is widely demonstrated, not only within Cisco but also through his role as a NetAcad instructor, where he has mentored and guided aspiring network professionals.

Bradley's dedication to excellence and his ability to simplify complex concepts have made him a respected figure in the networking community. His contributions continue to shape the future of networking, driving innovation and excellence in the industry. Bradley's refreshing no-nonsense approach to problem-solving has earned him a credible reputation among customers and peers alike.

Mohammad Khalil is an experienced service provider and enterprise expert, having worked in several service provider networks within the MENA region. Currently, he is a leader with the Cisco Competitive Win Center team covering enterprise architecture and working closely with Cisco sales/technical teams on designing their solutions.

Mohammad is a passionate networking expert, following technology trends and certifications development by writing several work labs and design scenarios. He was honored to be one of the SMEs for the CCIE Service Provider blueprint, SME for updated content of the CCIE Enterprise, and president of the Jordan IPv6 Council (part of the IPv6FORUM).

About the Technical Reviewer

Brad Edgeworth, CCIE No. 31574 (R&S and SP), is an SD-WAN technical solutions architect at Cisco Systems. Brad is a distinguished speaker at Cisco Live, where he has presented on various topics. Before joining Cisco, Brad worked as a network architect and consultant for various Fortune 500 companies. Brad's expertise is based on enterprise and service provider environments, with an emphasis on architectural and operational simplicity. Brad holds a bachelor of arts degree in computer systems management from St. Edward's University in Austin, Texas. Brad can be found on X (formerly Twitter) as @BradEdgeworth.

Dedications

To my beautiful wife, Evelina, who has supported me throughout the development of this book. You have challenged my assumptions, sharpened my ideas, and made some parts truly exceptional.

—*Bradley Riapolov*

I want to dedicate this book with affection to my wife, Qamar, and my precious children, Sireen and Saeed, who have supported me and motivated me to accomplish this effort. To my wise father, who believed in me and encouraged me all the way. Lastly, to my second home Estarta, which provided me with the support and environment to walk through.

—*Mohammad Khalil*

Acknowledgments

I would like to thank the Cisco Press team, especially James Manly and Christopher Cleveland, for their patience, guidance, and consideration.

I would like to thank the technical editor, colleague, and friend, Brad Edgeworth, for his time, technical expertise, and readiness to generously share knowledge without expecting anything in return.

I would like to thank Bikram Gandhok, Cisco Certification Program Manager, for driving Cisco Certifications to stay relevant and impactful.

I would like to thank Kent Dailey, my steadfast accomplice, who has navigated both challenges and victories with equal zeal.

I would like to thank my mentors, Marty Fierbaugh and Emerson Moura, for helping me gain insights that extend beyond my own perceptions.

Finally, I would like to thank my co-author, Mohammad, for sharing the burden of this challenging endeavor.

—Bradley Riapolov

I would like to massively thank our development editor, Chris Cleveland, for his guidance and the contribution he provided. James Manly's support was remarkable; a big thanks for his kindness.

I want to thank our technical editor, Brad, for his great insights, patience, and kindness for sharing his amazing practical experience and guiding us throughout the content development.

And I don't want to forget to thank my friend, Adeeb Al-Husseini, for his time sharing his practical experience and insights.

Finally, I would like to thank my partner, Bradley, for his seamless cooperation, contribution, and kindness to share his ideas and thoughts with me.

—Mohammad Khalil

Contents at a Glance

Online Elements

Contents

Online Elements

Command Syntax Conventions

The conventions used to present command syntax in this book are the same conventions used in the IOS Command Reference. The Command Reference describes these conventions as follows:

- **Boldface** indicates commands and keywords that are entered literally as shown. In actual configuration examples and output (not general command syntax), boldface indicates commands that are manually input by the user (such as a **show** command).
- *Italic* indicates arguments for which you supply actual values.
- Vertical bars (|) separate alternative, mutually exclusive elements.
- Square brackets ([ ]) indicate an optional element.
- Braces ({ }) indicate a required choice.
- Braces within brackets ([{ }]) indicate a required choice within an optional element.

Preface

I'm not an academic or engineer with a PhD, just an ordinary person like many of you. Back in 2000, when I lost my job, my college roommate's brother-in-law handed me a book on networking (*The CCNA Certification Guide*). I started my career at the very bottom, shlepping printers. I worked with various customers for eight years, thinking I was a pretty smart engineer until I joined Cisco, where smart engineers were as common as raindrops in a storm. Throughout my career, I've been fortunate to work with many networks, from small to massive, for over a quarter of a century. The insights in this book come from someone who's built real networks and learned from clever mentors, mistakes, and experiences.

While my writing style might not adhere to academic conventions, it remains practical, mirroring my approach to problem-solving. It offers an opportunity to assess your problem-solving approach and perhaps discover new perspectives. I aim to help you pass a particularly challenging exam and impart valuable skills to those who build networks, not just academics, whom I deeply respect. So, expect my approach to differ from what you're used to, with a focus on hands-on, real-world scenarios.

What makes this exam so tough? Having worked across various networking domains, I find that service provider networks are especially complex compared to their enterprise, data center, or mobility counterparts. Cisco acknowledges this complexity in its exam structure. Expect tough questions, and be pleasantly surprised when you encounter the ones easier than that.

As you prepare, pay close attention to the exam blueprint and how its sections are worded. Questions deliberately fall into four categories: Describe, Compare, Configure, and Troubleshoot. "Describe" sections assume a solid grasp of general knowledge and concepts. "Compare" sections probe deeper, expecting candidates to differentiate between similar topics. "Configure" sections require advanced familiarity to configure or spot errors accurately. "Troubleshoot" sections demand the highest skill level, testing your ability to solve complex problems with twists.

What's my top tip for acing the exam on your first attempt? Practice. I have structured my portions of the content for you to follow in the book and the command line. In theory, there is no difference between theory and practice. I can tell you that in practice, there is. There's no substitute for hands-on keyboard time in mastering service provider networks.

Once, at a networking convention, I spotted someone wearing a humorous T-shirt defining "engineer" as "someone who does precision guesswork based on unreliable data provided by those of questionable knowledge." During the exam and throughout your career, you might feel like that. But remember, you're not alone. This book is dedicated to those who've tackled the seemingly impossible in the past, those fixing networks alongside me today, and those learning to do the same in the near future. May some of us meet and recognize each other.

—Bradley Riapolov

To add to what Bradley mentioned, we tried our best to add new terms within the book from brainstorming and use cases to make it more realistic from practical experience and to assist with some design guidelines for the service provider technologies.

—Mohammad Khalil

Introduction

The Implementing and Operating Cisco Service Provider Network Core Technologies (SPCOR 350-501) exam is the required "core" exam for the CCNP Service Provider certifications. This exam tests your knowledge of implementing core service provider network technologies including architecture, services, networking, automation, quality of service, security, and network assurance.

Implementing and Operating Cisco Service Provider Network Core Technologies (SPCOR 350-501) is a 120-minute exam.

TIP You can review the exam blueprint from Cisco's website at https://learningnetwork.cisco.com/s/spcor-exam-topics.

This book gives you the foundation and covers the topics necessary to start your CCNP Service Provider or CCIE Service Provider journey.

The CCNP Service Provider Certification

The CCNP Service Provider certification is one of the industry's most respected certifications. In order for you to earn the CCNP Service Provider certification, you must pass two exams: the SPCOR exam covered in this book (which covers core security technologies) and one of four available service provider concentration exams of your choice, so you can customize your certification to your technical area of focus.

TIP The SPCOR core exam is also the qualifying exam for the CCIE Service Provider certification. Passing this exam is the first step toward earning both of these certifications.

The following are the CCNP Service Provider concentration exams:

- Implementing Cisco Service Provider Advanced Routing Solutions (300-510 SPRI)
- Implementing Cisco Service Provider VPN Services (300-515 SPVI)
- Automating Cisco Service Provider Solutions (300-535 SPAUTO)
- Designing and Implementing Cisco Service Provider Cloud Network Infrastructure (300-540 SPCNI)

The CCIE Service Provider Certification

The CCIE Service Provider certification is one of the most admired, elite, and challenging certifications in the industry. The CCIE Service Provider program prepares you to be a recognized technical leader. In order to earn the CCIE Service Provider certification, you must pass the SPCOR 350-501 exam and an eight-hour, hands-on lab exam. The lab exam covers very complex network service provider network scenarios. These scenarios range from designing through implementing, operating, and optimizing dual-stack solutions (IPv4 and IPv6) of complex service provider networks.

Cisco considers ideal candidates to be those who possess the following:

- Extensive hands-on experience with Cisco's Service Provider portfolio
- Experience deploying Cisco's wide assortment of legacy and modern service provider technologies
- Deep understanding of multiple transport protocols and multitenant segmentation solutions
- Hands-on experience with MPLS networks and VPN solutions
- Configuring and troubleshooting QoS, mobility networking, device hardening, and general and access control
- Deep understanding of network automation and orchestration constructs

The Exam Objectives (Domains)

The Implementing and Operating Cisco Service Provider Network Core Technologies (SPCOR 350-501) exam is broken down into five major domains. The contents of this book cover each of the domains and the subtopics included in them, as illustrated in the following descriptions.

The following table breaks down each of the domains represented in the exam.

Domain	Percentage of Representation in Exam
1: Architecture	15%
2: Networking	30%
3: MPLS and Segment Routing	20%
4: Services	20%
5: Automation and Assurance	15%
	Total 100%

Here are the details of each domain:

Domain 1: Architecture: This domain is covered in Chapters 1–3, 14, 16–18, and 21.

1.1 Describe service provider architectures

1.1.a Core architectures (Metro Ethernet, MPLS, unified MPLS, SR, SRTE, SRv6)

1.1.b Transport technologies (xDSL, DWDM, DOCSIS, TDM, and xPON)

1.1.c Mobility (packet core, RAN xhaul transport for 5G vRAN and ORAN transport)

1.1.d Routed optical network

1.2 Describe Cisco network software architecture

1.2.a IOS

1.2.b IOS XE

1.2.c IOS XR

1.3 Describe service provider virtualization

1.3.a NFV infrastructure

1.3.b VNF workloads

1.3.c Containers

1.3.d Application hosting

1.4 Describe QoS architecture

1.4.a MPLS QoS models (Pipe, Short Pipe, and Uniform)

1.4.b MPLS TE QoS (MAM, RDM, CBTS, PBTS, and DS-TE)

1.4.c DiffServ and IntServ QoS models

1.4.d Trust boundaries between enterprise and SP environments

1.4.e IPv6 flow label

1.5 Configure and verify control plane security

1.5.a Control plane protection techniques (LPTS and CoPP)

1.5.b BGP-TTL security and protocol authentication

1.5.c BGP prefix suppression

1.5.d LDP security (authentication and label allocation filtering)

1.5.e BGP sec

1.5.f BGP flowspec

1.6 Describe management plane security

1.6.a Traceback

1.6.b AAA and TACACS

1.6.c RestAPI security

1.6.d DDoS

1.7 Implement data plane security

1.7.a uRPF

1.7.b ACLs

1.7.c RTBH

1.7.d MACsec

Domain 2: Networking: This domain is covered in Chapters 4–8, 19, and 20.

2.1 Implement IS-IS (IPv4 and IPv6)

2.1.a Route advertisement

2.1.b Area addressing

2.1.c Single/Multitopology

2.1.d Metrics

2.2 Implement OSPF (v2 and v3)

2.2.a Neighbor adjacency

2.2.b Route advertisement

- 2.2.c Multiarea (addressing and types)
- 2.2.d Metrics

2.3 Describe BGP path selection algorithm

2.4 Implement BGP (v4 and v6 for IBGP and EBGP)

- 2.4.a Neighbors
- 2.4.b Prefix advertisement
- 2.4.c Address family
- 2.4.d Path selection
- 2.4.e Attributes
- 2.4.f Redistribution
- 2.4.g Additional Paths
- 2.4.h PIC

2.5 Implement routing policy language and route maps (BGP, OSPF, IS-IS)

2.6 Troubleshoot routing protocols

- 2.6.a Neighbor adjacency (BGP, OSPF, IS-IS)
- 2.6.b Route advertisement (BGP, OSPF, IS-IS)

2.7 Describe IPv6 transition (NAT44, NAT64, CGNAT, MAP-T and DS Lite)

2.8 Implement high availability

- 2.8.a NSF / graceful restart
- 2.8.b NSR
- 2.8.c BFD
- 2.8.d Link aggregation

Domain 3: MPLS and Segment Routing: This domain is covered in Chapters 10, 14, and 15.

3.1 Implement MPLS

- 3.1.a LDP sync
- 3.1.b LDP session protection
- 3.1.c LDP neighbors
- 3.1.d Unified MPLS
- 3.1.e MPLS OAM

3.2 Describe traffic engineering

- 3.2.a IS-IS and OSPF extensions
- 3.2.b RSVP functionality
- 3.2.c FRR

3.3 Describe segment routing

- 3.3.a Segment types
- 3.3.b SR control plane (BGP, OSPF, IS-IS)
- 3.3.c Segment routing traffic engineering

3.3.d TI-LFa

3.3.e PCE-PCC architectures

3.3.f Flexible algorithm

3.3.g SRv6 (locator, micro-segment, encapsulation, interworking gateway)

Domain 4: Services: This domain is covered in Chapters 9, 11–13, and 21.

4.1 Describe VPN services

4.1.a EVPN

4.1.b Inter-AS VPN

4.1.c CSC

4.1.d mVPN

4.2 Configure L2VPN and Carrier Ethernet

4.2.a Ethernet services (E-Line, E-Tree, E-Access, E-LAN)

4.2.b IEEE 802.1ad, IEEE 802.1ah, and ITU G.8032

4.2.c Ethernet OAM

4.2.d VLAN tag manipulation

4.3 Configure L3VPN

4.3.a Intra-AS VPN

4.3.b Shared services (extranet and Internet)

4.4 Implement multicast services

4.4.a PIM (PIM-SM, PIM-SSM, and PIM-BIDIR, PIMv6)

4.4.b IGMP v1/v2/v3 and MLD

4.5 Implement QoS services

4.5.a Classification and marking

4.5.b Congestion avoidance, traffic policing, and shaping

Domain 5: Automation and Assurance: This domain is covered in Chapter 22.

5.1 Describe the programmable APIs used to include Cisco devices in network automation

5.2 Interpret an external script to configure a Cisco device using a REST API

5.3 Describe the role of Network Services Orchestration (NSO)

5.4 Describe the high-level principles and benefits of a data modeling language, such as YANG

5.5 Describe configuration management tools, such as Ansible and Terraform

5.6 Describe Secure ZTP

5.7 Configure dial-in/out, TCP, TLS, and mTLS certificates using gRPC and gNMI

5.8 Configure and verify NetFlow/IPFIX

5.9 Configure and verify NETCONF and RESTCONF

5.10 Configure and verify SNMP (v2c/v3)

Steps to Pass the SPCOR Exam

There are no prerequisites for the SPCOR exam. However, students must have an understanding of networking and cybersecurity concepts.

Signing Up for the Exam

The steps required to sign up for the Implementing and Operating Cisco Service Provider Network Core Technologies (SPCOR 350-501) exam:

1. Create a Certiport account at https://www.certiport.com/portal/SSL/Login.aspx.
2. Once you have logged in, make sure that "Test Candidate" from the drop-down menu is selected.
3. Click the **Shop Available Exams** button.
4. Select the **Schedule exam** button under the exam you wish to take.
5. Verify your information and continue throughout the next few screens.
6. On the **Enter payment and billing** page, click the **Add Voucher or Promo Code** button if applicable. Enter the voucher number or promo/discount code in the field below and click the **Apply** button.
7. Continue through the next two screens to finish scheduling your exam.

Facts About the Exam

The exam is a computer-based test. The exam consists of multiple-choice questions only. You must bring a government-issued identification card. No other forms of ID will be accepted. You can take the exam at a Pearson Vue center or online via the OnVUE platform. Visit the OnVUE page for your exam program: https://home.pearsonvue.com/Test-takers/OnVUE-online-proctoring/View-all.aspx

Once there, navigate to the FAQs section of the page, where you'll find helpful information on everything from scheduling your exam to system requirements, testing policies, and more.

NOTE Refer to the Cisco Certification site at https://cisco.com/go/certifications for more information regarding this, and other, Cisco certifications.

About the *CCNP SPCOR 350-501 Official Cert Guide*

This book maps directly to the topic areas of the SPCOR exam and uses a number of features to help you understand the topics and prepare for the exam.

Objectives and Methods

This book uses several key methodologies to help you discover the exam topics that need more review, to help you fully understand and remember those details, and to help you prove to yourself that you have retained your knowledge of those topics. The book does not try to help you pass the exam only by memorization; it seeks to help you to truly learn and understand the topics. This book is designed to help you pass the Implementing and Operating Cisco Service Provider Network Core Technologies (SPCOR 350-501) exam by using the following methods:

- Helping you discover which exam topics you have not mastered
- Providing explanations and information to fill in your knowledge gaps
- Supplying exercises that enhance your ability to recall and deduce the answers to test questions
- Providing practice exercises on the topics and the testing process via test questions on the companion website

How to Use This Book

To help you customize your study time using this book, the core chapters have several features that help you make the best use of your time:

- **Foundation Topics:** These are the core sections of each chapter. They explain the concepts for the topics in that chapter.
- **Brainstorming:** These encourage you to actively apply their knowledge. You are invited to take a step beyond mere fact retrieval and engage deeply with the material you've covered. This is your opportunity to think critically about what you've learned and attempt to apply it on your own. Most of the time, we guide you through the process, but these exercises are designed to help you assess an exam question and provide an educated, well-reasoned answer. By participating in these sessions, you'll develop the skills to approach challenges with confidence and creativity, ensuring you're prepared to succeed independently.
- **Exam Preparation Tasks:** After the "Foundation Topics" section of each chapter, the "Exam Preparation Tasks" section lists a series of study activities that you should do at the end of the chapter:
  - **Review All Key Topics:** The Key Topic icon appears next to the most important items in the "Foundation Topics" section of the chapter. The Review All Key Topics activity lists the key topics from the chapter, along with their page numbers. Although the contents of the entire chapter could be on the exam, you should definitely know the information listed in each key topic, so you should review these.
  - **Define Key Terms:** Although the Implementing and Operating Cisco Service Provider Network Core Technologies (SPCOR 350-501) exam may be unlikely to ask a question such as "Define this term," the exam does require that you learn and know a lot of cybersecurity terminology. This section lists the most important terms from the chapter, asking you to write a short definition and compare your answer to the glossary at the end of the book.

  - **Review Questions:** Confirm that you understand the content you just covered by answering these questions and reading the answer explanations in Appendix A.

- **Web-based practice exam:** The companion website includes the Pearson Cert Practice Test engine, which allows you to take practice exam questions. Use it to prepare with a sample exam and to pinpoint topics where you need more study.

How This Book Is Organized

This book contains 22 core chapters—Chapters 1 through 22. Chapter 23 includes preparation tips and suggestions for how to approach the exam. Each core chapter covers a subset of the topics on the Implementing and Operating Cisco Service Provider Network Core Technologies (SPCOR 350-501) exam. The core chapters map to the SPCOR topic areas and cover the concepts and technologies you will encounter on the exam.

The Companion Website for Online Content Review

All the electronic review elements, as well as other electronic components of the book, exist on this book's companion website.

To access the companion website, which gives you access to the electronic content that accompanies this book, start by establishing a login at www.ciscopress.com and registering your book by December 31, 2027. To do so, simply go to www.ciscopress.com/register and enter the ISBN of the print book: 9780135324806. After you have registered your book, go to your account page and click the **Registered Products** tab. From there, click the **Access Bonus Content** link to get access to the book's companion website.

Note that if you buy the Premium Edition eBook and Practice Test version of this book from Cisco Press, your book will automatically be registered on your account page. Simply go to your account page, click the **Registered Products** tab, and select **Access Bonus Content** to access the book's companion website.

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

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.

How to Access the Pearson Test Prep (PTP) App

You have two options for installing and using the Pearson Test Prep application: a web app and a desktop app. To use the Pearson Test Prep application, start by accessing the registration code that comes with the book. You can access the code in these ways:

- You can get your access code by registering the print ISBN 9780135324806 on https://www.ciscopress.com/register. Make sure to use the print book ISBN, regardless of whether you purchased an eBook or the print book. After you register the book, your access code will be populated on your account page under the Registered Products tab. Instructions for how to redeem the code are available on the book's companion website by clicking the **Access Bonus Content** link.

- If you purchase the Premium Edition eBook and Practice Test directly from the Pearson IT Certification website, the code will be populated on your account page after purchase. Just log in at https://www.ciscopress.com, click **Account** to see details of your account, and click the **Digital Purchases** tab.

NOTE After you register your book, your code can always be found in your account under the Registered Products tab.

Once you have the access code, to find instructions about both the PTP web app and the desktop app, follow these steps:

Step 1. Open this book's companion website, as was shown earlier in this Introduction under the heading "The Companion Website for Online Content Review."

Step 2. Click the **Practice Exams** button.

Step 3. Follow the instructions listed there both for installing the desktop app and for using the web app.

Note that if you want to use the web app only at this point, just navigate to www.pearsontestprep.com, establish a free login if you do not already have one, and register this book's practice tests using the registration code you just found. The process should take only a couple of minutes.

Customizing Your Exams

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

- **Study mode:** Allows you to fully customize your exams and review answers as you are taking the exam. This is typically the mode you would use first to assess your knowledge and identify information gaps.
- **Practice Exam mode:** Locks certain customization options, as it is presenting 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 do, so you should not use it if you are trying to 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 well as 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.

There are several other customizations you can make to your exam from the exam settings screen, such as the time of the exam, the number of questions served up, whether to randomize questions and answers, whether to show the number of correct answers for multiple-answer questions, and 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 while connected to the Internet, it checks if there are any updates to your exam data and automatically downloads any changes that were made since the last time you used the software.

Sometimes, due to many factors, the exam data may not fully download when you activate your exam. If you find that figures or exhibits are missing, you may need to manually update your exams. To update a particular exam you have already activated and downloaded, simply click the **Tools** tab and click the **Update Products** button. Again, this is only an issue with the desktop Windows application.

If you wish to check for updates to the Pearson Test Prep exam engine software, Windows desktop version, simply click the **Tools** tab and click the **Update Application** button. This ensures that you are running the latest version of the software engine.

Figure Credits

Figure 1-15 © ITU 2024

Figure 7-2, baranoski.ca

Figures 17-1 and 17-2 © 2024 Postman, Inc.

Cover, insta_photos/shutterstock

CHAPTER 15

Segment Routing

This chapter covers the following exam topics:

3.3 Describe segment routing

- 3.3.a Segment types
- 3.3.b SR control plane (BGP, OSPF, IS-IS)
- 3.3.c Segment routing traffic engineering
- 3.3.d TI-LFA
- 3.3.e PCE-PCC architectures
- 3.3.f Flexible algorithm
- 3.3.g SRv6 (locator, micro-segment, encapsulation, interworking gateway)

Segment Routing represents a cutting-edge technology designed to enhance and optimize the use of MPLS-based and IPv6 networks. This innovative approach introduces a suite of tools and concepts that not only simplify network operations but also significantly enhance the flexibility available to network operators. By allowing for more granular control over packet paths, Segment Routing empowers network operators to efficiently navigate the intricacies of modern networks, adapt to dynamic requirements, and minimize the challenges associated with traditional MPLS-based architectures.

In essence, Segment Routing is a beacon of progress in the networking realm, offering a robust set of solutions to meet the most demanding network requirements. Its advent marks a transformative shift in how networks are managed, providing operators with a versatile toolkit to navigate the complexities of traffic engineering, and ensuring that networks can dynamically respond to changing conditions in real time. By simplifying the routing paradigm and enhancing the scalability of MPLS-based networks, Segment Routing emerges as a pivotal technology, poised to play a crucial role in the evolution of modern network architectures.

“Do I Know This Already?” Quiz

The “Do I Know This Already?” quiz allows you to assess whether you should read this entire chapter thoroughly or jump to the “Exam Preparation Tasks” section. If you are in doubt about your answers to these questions or your own assessment of your knowledge of the topics, read the entire chapter. Table 15-1 lists the major headings in this chapter and their corresponding “Do I Know This Already?” quiz questions. You can find the answers in Appendix A, “Answers to the ‘Do I Know This Already?’ Quizzes and Review Questions.”

Table 15-1 "Do I Know This Already?" Section-to-Question Mapping

Foundation Topics Section	Questions
Segment Types	1–2
Segment Routing Control Plane	3–4
Segment Routing Traffic Engineering	5–8
PCE-PCC Architecture	9–10

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

1. Which of the following are viable options for assigning SRGB ranges? (Select all that apply.)
   - **a.** Globally
   - **b.** Per IGP
   - **c.** Dynamically
   - **d.** Statically
2. Which labels cannot be a part of SRGB? (Select all that apply.)
   - **a.** 15999
   - **b.** 16000
   - **c.** 24999
   - **d.** 1,048,576
3. Which TLVs play an important role in Segment Routing? (Select all that apply.)
   - **a.** 22
   - **b.** 41
   - **c.** 135
   - **d.** 163
4. Which of the following is not an option for the SR Control Plane?
   - **a.** RSVP-TE
   - **b.** IS-IS
   - **c.** OSPF
   - **d.** BGP
5. Which of the following is not a viable option to supply SR policy to the ingress router?
   - **a.** NETCONF
   - **b.** FIB

  - **c.** CLI
  - **d.** PCEP

6. Which behavior properly describes an SR policy?

  - **a.** An SR policy uses Network Service Orchestrator (NSO) to program its data plane.
  - **b.** SR-PCE will use BGP-LS to program an SR policy into the ingress node.
  - **c.** An SR policy encodes the list of constraints into the MPLS headers.
  - **d.** An SR policy will use BSID entries to program its forwarding table.

7. Which of the following describes TI-LFA functionality? (Select all that apply.)

  - **a.** TI-LFA must rely on LDP functionality to provide double-segment protection.
  - **b.** TI-LFA uses next-hop neighbor as the point of local repair.
  - **c.** TI-LFA preprograms the post-convergence path into a router's data plane.
  - **d.** TI-LFA can use P space to calculate the post-convergence path.

8. Viable repair tunnel endpoints are found at which intersections?

  - **a.** At midpoints of extended spaces
  - **b.** At the intersection of point of local repair and double segments
  - **c.** At the intersection of PQ nodes
  - **d.** At the intersection of P and Q spaces

9. Which component serves the needs of long-term network engineering and capacity planning?

  - **a.** Crosswork Network Controller
  - **b.** Crosswork Optimization Engine
  - **c.** WAN Automation Engine
  - **d.** Crosswork Cloud

10. Which of the following protocols are used by PCEP to calculate network paths? (Select all that apply.)

  - **a.** IS-IS
  - **b.** BGP-LS
  - **c.** RSVP-TE
  - **d.** OSPF

Foundation Topics

Over time, a skilled engineer learns to make the most of the available resources. Over the past three decades, various strategies have been explored to maximize the potential of computer networks. The traditional approach to steering network traffic has been to manipulate the Interior Gateway Protocol (IGP) and rely solely on destination-based routing strategies. This only provided a restricted range of options for directing traffic within the network. The main limitation of using IGP metrics has been the "all or nothing" approach. Some of the links inevitably become congested while other links remain underutilized even though they are high bandwidth or low latency. IGP metrics simply lack optimization capabilities because they do not allow you to map different services to different paths. Despite years of featured improvements, such challenges persist and remain unsolvable with conventional IGP manipulations and destination-based routing strategies.

Thus, in the 1990s, the development of the **Label Distribution Protocol (LDP)** and Resource Reservation Protocol with Traffic Engineering extensions (RSVP-TE) marked a significant advancement in networking. These capable adjunct protocols effectively addressed specific challenges and provided network operators with crucial Traffic Engineering tools to overcome a wide variety of traffic-steering issues. Network operators could now manipulate how traffic flowed through the network to make the most of the available paths. Nevertheless, while providing network optimization, these protocols also introduced intricate challenges.

In the case of LDP, an additional process had to be created and maintained on the network, which led to a complex interaction with the Interior Gateway Protocol. LDP-IGP synchronization problems would cause traffic disruptions until these two protocols could settle on an agreement regarding the best way to forward traffic.

When dealing with RSVP-TE, reserving bandwidth accurately involves placing traffic within RSVP-TE tunnels. While this approach is feasible in smaller networks with minimal traffic engineering needs, it becomes exceedingly intricate on a larger scale. Managing hundreds of tunnels, their backup paths, and upkeep of a pertinent set of rules in the face of a dynamically evolving network posed formidable challenges, demanding considerable time and effort. Such scaling issues led many operators to limit their RSVP-TE deployments to the fast reroute (FRR) use cases. RSVP-TE is also not “ECMP-friendly,” so it can never use all IGP-derived paths, forcing the operator to create more tunnels. An additional aspect to consider is that RSVP-TE generates a persistent “always-on” network state where every router must account for available bandwidth and that state must be constantly monitored. This incurs a cost in terms of network compute resources and the hardware required to sustain this continuous state irrespective of whether the network experiences congestion or not.

Could network optimization be further improved? These were the experiences and thoughts of the designers behind Segment Routing. They proposed that a properly designed network should have enough capacity to effectively handle an expected volume of traffic without congestion, even in the presence of a probable set of independent failures. IGP coupled with ECMP can competently absorb the majority of the traffic volume. In less frequent instances of congestion, traffic engineering tools would address applications intolerant of such network bottlenecks. This represented a simpler and more resource-efficient approach, both in terms of hardware and human efforts.

They proposed that it is better to distribute labels associated with IGP-signaled prefixes within the IGP framework itself, rather than relying on a separate protocol such as LDP to perform this task. This would solve the LDP-IGP synchronization problem during network failures because there now would be a single source of truth (IGP) to find available network paths. The network would precalculate such paths even before the failure occurred. IGP can have such “preknowledge”—think of an EIGRP-feasible successor—which is aware of the best available network route even before the failure occurs.

Second, why not give the network operator the power to direct any packet to any path at the ingress router only (where the packet enters the network), without having to maintain the expensive and complex “always-on” state throughout the entire network domain? This would lower control plane pressure, conserve network resources, and give the operator the full flexibility to force any packet anywhere.

Cisco engineers formulated the Segment Routing concept, obtained approval from their management in 2012, presented the idea to IETF in March 2013, authored numerous IETF drafts, and significantly influenced the industry in 2015 with the introduction of Segment Routing on IOS XR platforms (IOS and NX-OS carry some of the SR features). The market positively responded with other vendors supporting this capable technology. It is also referred to as Source Routing outside of Cisco. As of the present moment, Cisco alone has documented more than 1200 operational Segment Routing production deployments.

Segment Routing can be deployed on two data planes:

- MPLS data plane where segments will be encoded with MPLS labels.
- IPv6 data plane where segments will be encoded with IPv6 addresses.

Segment Types

Think of segments as a set of instructions. In Segment Routing, a source (an ingress router, as an example) chooses a certain path through the network and encodes the path in the packet header as an ordered list of instructions. These instructions are termed *segments* because they describe components of a divided whole.

In **SR-MPLS (Segment Routing based on MPLS data plane)**, such an identifier refers to an ordered list of segments represented by a stack of MPLS labels. When you instruct routers to follow these labels in a given sequence, the packets will take this path through the network. In **SRv6 (Segment Routing based on IPv6 data plane)**, it refers to an ordered list of segments encoded into a routing extension header. When you instruct routers to follow this list, the packets will flow via this path. In Figure 15-1, assigning such identifiers to routers can provide instructions for a specific path to send packets.

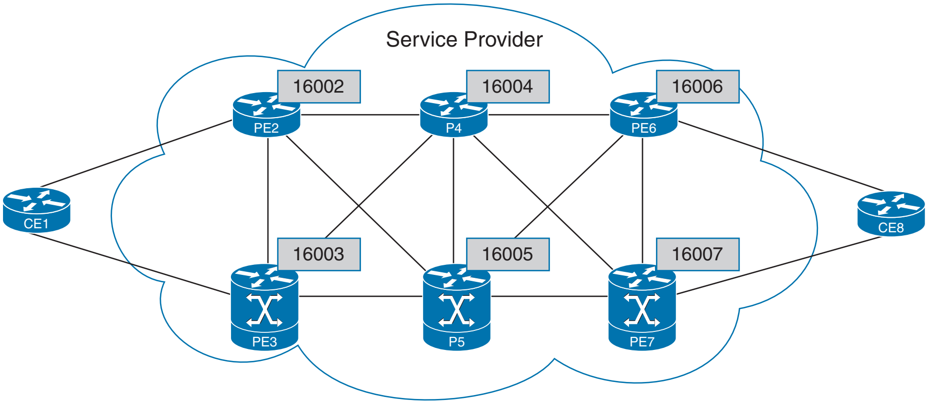


Figure 15-1 *Assigning Segments to Network*

Global Segments

Every router in a Segment Routing, or SR, domain understands such instruction and installs it in its forwarding table. This instruction is a domain-wide (watch the misleading name *global* because it is not known globally around the world) unique label value (a numerical number) that comes from the **Segment Routing Global Block (SRGB)** database. Table 15-2 shows how the **Label Switching Database (LSD)** carves the following default ranges (some can be changed) on Cisco routers running Segment Routing–capable software.

Key Topic

Table 15-2 LSD Label Ranges

Label Range	Reserved for	Examples
0–15	Base special-purpose MPLS labels	0—IPv4 Explicit NULL 3—Implicit NULL
16–15,999	Static MPLS labels	LDP assigned
16,000–23,999	SRGB	Global Segments (i.e., SR)
24,000–1,048,575	Dynamic allocation	Adjacency Segments

In Figure 15-1, every node in the domain knows that label 16002 always and uniquely represents Router PE2, 16003 always represents Router PE3, and so on. These are referred to as Node SIDs (Segment Identifiers).

A Node SID is a type of Prefix SID, as it represents any prefix linked to a node. To send a set of segment routing instructions is to specify these labels (Node SIDs); 16002, 16003 literally means "send this traffic via shortest ECMP path to PE2, then same to PE3." If the "uniqueness" rule is broken (two distinct routers are assigned the same label value), there will be an issue on your network because the nodes will not be able to accurately determine the appropriate path for routing traffic. (Technically, what happens is that IS-IS will prioritize the "first programmed" label and ignore the "second" one. This can get complex as to what "first programmed" means, as in when a router reboots or has failed, which router becomes "first programmed"? OSPF, on the other hand, will withdraw both SIDs. Don't worry about this because this topic is highly unlikely to show up on the exam, but good to know.)

Global Segments are always distributed as a unique value via IGP (remember that SR no longer relies on LDP as the label distribution mechanism). This value must be unique and comes from a combination of a label range (SRGB) + index. The default SRGB range for Cisco routers is 16000–23999 (notice how it does not overlap with the LDP range) and the index is zero based (the first index = 0). In our scenario, we start adding index values to our routers (index 2 for PE2, index 3 for PE3), so we will come up with globally known unique values of 16002 and 16003 for these routers. Cisco gives an option to also define these as an absolute value. There is no difference in daily operations whether absolute or relative indexes are used.

There are two minimum requirements to enable Segment Routing:

- Configure the Segment Routing Global Block (SRGB)
- Enable Segment Routing and Node SID in the IGP (shown later in the chapter under respective protocols)

Example 15-1 demonstrates SRGB assignment on both IOS XE as well as IOS XR operating systems.

Example 15-1 *SRGB Verification on IOS XE and IOS XR*

```
IOS
PE2# show running-config segment routing
!
```

```
segment-routing
 global-block 16000 23999
!
PE2#
```

```
IOS XR
RP/0/0/CPU0:PE4# show running-config segment-routing
!
segment-routing
 global-block 16000 23999
!
RP/0/0/CPU0:PE4#
```

When you're configuring an SRGB block, the recommendation is to configure it globally (again misleading, in the global, i.e., router-wide configuration), not per individual IGP; this way, all IGP instances as well as BGP can use the global (i.e., router-wide) SRGB. This is important because, later on, if you choose to add another protocol and the SRGB range has to change, you will have to reload the router. To avoid this, it is better to assign the SRGB block globally in the router so that multiple protocols (including BGP) can use it, not just a specific IGP. You should have a homogenous SRGB block (same SRGB range) on all routers on the network. If you do not, you will have fun times building end-to-end LSPs.

Local Segments

This instruction is allocated and understood by the originating node. It is locally significant only (which aligns with the intended meaning regarding the local router). A locally allocated MPLS label would be a good example of a **local segment**. Sometimes a router has more than a single link for forwarding traffic. The operator can prefer one of these paths over the other. In Figure 15-2, a packet arrives at PE3 (via label 16003). PE3 intends to send this packet to PE6, and there are two shortest available paths—through P4 and P5. These links are identified with local labels 24100 and 24150. Selecting one of them will instruct the local router to pick the appropriate link.

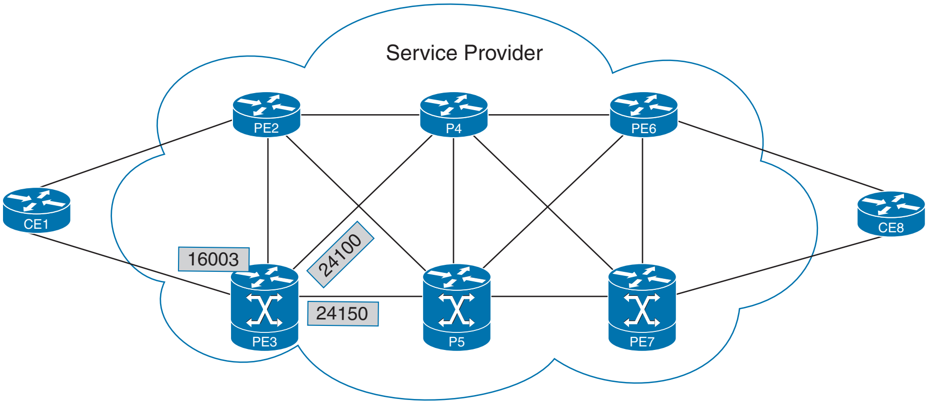


Figure 15-2 *Assigning Local Segments to Network*

IGP Segments

Segments construct a path through a network. There are two building blocks distributed by IGP: Prefix Segments and Adjacency Segments.

IGP Prefix Segments

Think of an **IGP Prefix Segment** as a shortest path to the IGP prefix. Note that it is

- Known as Prefix-SID
- Associated with an IP prefix
- Represents an ECMP (Equal Cost Multi-Path)-aware shortest path to a prefix
- Likely a multihop path
- A Global Segment—known uniquely on the SR domain
- A Label (16000 + Index) that is advertised as an index
- Distributed via IGP (ISIS/OSPF)

Observe this behavior in Figure 15-3 (note that we removed two links to better illustrate our example), where different routers are used to forward traffic to P5 based on label 16005, destined to the loopback address of 10.1.100.5. For this Segment Routing domain, PE3, P4, PE6, and PE7 send traffic directly to P5 because when these routers look at the MPLS forwarding table, label 16005 will be associated with the one interface directly connecting to P5. The only exception is PE2 which will have two equal paths available due to ECMP. How did all the routers learn that prefix 10.1.100.5/32 is associated with label 16005? R5 has generated this value from the combination of the SRGB label base 16000, added its operator assigned index +5, and advertised this label via IGP.

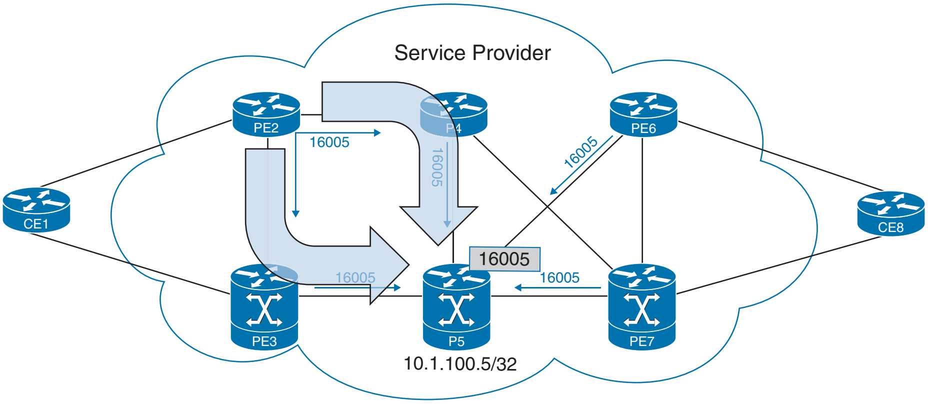


Figure 15-3 *IGP Prefix Segment Behavior*

I need to point something out here. Technically, there is a difference between a Node SID (which we just showed you) and a Prefix SID. A Node SID points to a router that can be a visiting point on the network. A Prefix SID is a label for a network prefix that is advertised by a router. This causes a point of confusion at times. Note that a special N flag is set to indicate that a SID represents a router (node) on the network. This is advanced and is unlikely to show up on the exam, but we include Example 15-2 to clear up any confusion regarding the difference.

Example 15-2 *Node SID and the N Flag*

```
RP/0/0/CPU0:R2# show isis database R1.00-00 detail verbose
Thu May  9 13:03:44.757 UTC
IS-IS lab (Level-1) Link State Database
LSPID                 LSP Seq Num  LSP Checksum  LSP Holdtime/Rcvd  ATT/P/OL
The requested LSP R1.00-00 was not found in the IS-IS lab Level-1 LSP Database
IS-IS lab (Level-2) Link State Database
LSPID                 LSP Seq Num  LSP Checksum  LSP Holdtime/Rcvd  ATT/P/OL
R1.00-00              0x00000007   0xf6b7        1108 /1199         0/0/0
  Area Address:   49.0001
  NLPID:          0xcc
  NLPID:          0x8e
  IP Address:     1.1.1.1
  Metric: 10         IP-Extended 10.1.12.0/24
    Prefix Attribute Flags: X:0 R:0 N:0
  Metric: 10         IP-Extended 10.1.13.0/24
    Prefix Attribute Flags: X:0 R:0 N:0
  Metric: 0          IP-Extended 1.1.1.1/32
    Prefix-SID Index: 1, Algorithm:0, R:0 N:1 P:0 E:0 V:0 L:0
    Prefix Attribute Flags: X:0 R:0 N:1
  Hostname:       R1
  IPv6 Address:   2001:1:1:1::1
  Metric: 10         MT (IPv6 Unicast) IPv6 2001:10:1:12::/64
    Prefix Attribute Flags: X:0 R:0 N:0
  Metric: 10         MT (IPv6 Unicast) IPv6 2001:10:1:13::/64
    Prefix Attribute Flags: X:0 R:0 N:0
  Metric: 0          MT (IPv6 Unicast) IPv6 2001:1:1:1::1/128
    Prefix-SID Index: 1001, Algorithm:0, R:0 N:1 P:0 E:0 V:0 L:0
    Prefix Attribute Flags: X:0 R:0 N:1
  MT:             Standard (IPv4 Unicast)
  MT:             IPv6 Unicast                                 0/0/0
  Metric: 10         IS-Extended R2.07
    Interface IP Address: 10.1.12.1
    Link Maximum SID Depth:
      Label Imposition: 10
    LAN-ADJ-SID: F:0 B:0 V:1 L:1 S:0 P:0 weight:0 Adjacency-sid: 24001 System ID:R2
  Metric: 10         MT (IPv6 Unicast) IS-Extended R2.07
    Interface IPv6 Address: 2001:10:1:12::1
Link Maximum SID Depth:
      Label Imposition: 10
    LAN-ADJ-SID: F:1 B:0 V:1 L:1 S:0 P:0 weight:0 Adjacency-sid: 24003 System ID:R2
  Router Cap:     1.1.1.1 D:0 S:0
```

```
    Segment Routing: I:1 V:1, SRGB Base: 16000 Range: 8000
    SR Local Block: Base: 15000 Range: 1000
    SR Algorithm:
      Algorithm: 0
      Algorithm: 1
    Node Maximum SID Depth:
      Label Imposition: 10
RP/0/0/CPU0:R2#
```

This example from IS-IS shows the R2 is receiving SR information from R1, which has IPv4 (1.1.1.1/32) and IPv6 (2001:1:1:1::1/128) prefixes advertised with the N (Node) flag set to 1—a Node SID, not a Prefix SID.

IGP Adjacency Segment

An **IGP Adjacency Segment** is an identifier that describes a particular link between two routers. It is used to direct traffic over a specific link within the IGP routing domain. Routers can be given an instruction to forward based on the IGP adjacency. Note that adjacency segment is

- Known as Adj-SID
- Represents a hop over a specific link between two IGP-speaking routers
- Likely a one-hop path
- A Local Segment—significant only on a particular router
- Advertised as a label value
- Distributed via IGP (ISIS/OSPF)

Figure 15-4 illustrates this. Now, the packet has arrived at P5 and needs to travel further to PE6 (10.1.100.6). The operator has a choice to impose a local decision on P5 on which links to use—the direct link to PE6 or two-hop link via PE7. In this case, the link to PE7 is preferred (higher bandwidth, lower latency, encryption—take your pick), and label 24150 steers the traffic toward PE7.

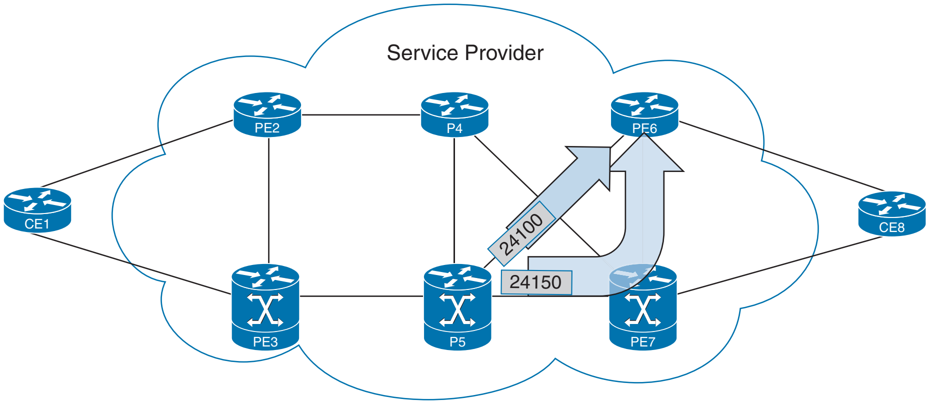


Figure 15-4 *IGP Prefix Segment Behavior with Adjacency Segments*

Combining IGP Segments

By combining segment IDs, you can groom traffic on any path in the network:

1. Specify the sequential list of segment IDs in the packet header, known as a label stack with the top label being read first.
2. Path is *not* signaled, and per flow state is not created (as in RSVP-TE).
3. A single protocol (IS-IS, OSPF, BGP) distributes this instruction.

In Figure 15-5, a network operator instructs PE2 to steer traffic to PE6 by sending it to P5 first via two available ECMP paths. Once P5 gets the packet, the top label 16005 is removed, and P5 uses the direct link to PE6 by looking up this interface in the forwarding table where it is associated with the dynamic adjacency label 24100.

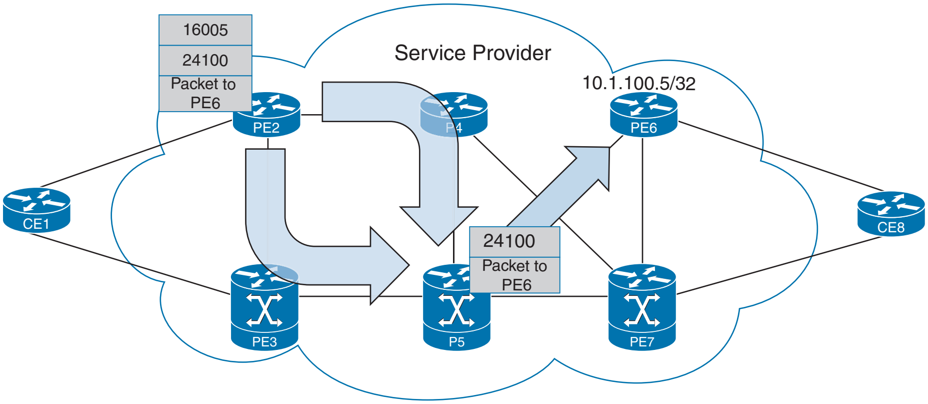


Figure 15-5 *Combining IGP Segment IDs for Traffic Steering*

By combining segment IDs (Prefix-SID and Adj-SID) in this way, you can put a packet on any path through the network, no matter how complex or unnatural this path may be. That is the power and essence of Segment Routing. At each hop, the top segment identifies the next hop. Segment IDs are stacked in sequential order at the top of the packet header. When the top segment ID contains the identity of another router, the receiving node uses equal cost multipaths (ECMP) to move the packet to the next hop. When the identity is the receiving router itself, the router will pop the top segment and perform the task required by the next segment.

NOTE Please note that I (Brad) am simplifying the mechanics of Segment Routing to its basic elements to facilitate a clear understanding of the fundamental concepts. In practice, multiple labels may be used, including transport labels and service labels that carry L3VPN traffic along with traffic engineering. Additionally, the number of labels a platform can handle depends on its hardware capabilities. Generally, Segment Routing can accommodate multiple labels in the label-switched path (LSP), with some platforms supporting up to 6, 9, or even 12 labels. However, most networks do not typically construct such elaborate paths, although it is possible and some customers have implemented them. Figure 15-6 shows a simple example of a label stack where PE2 assigns the inner service label 24192 for VPN traffic between CE1 and CE8. Labels are disposed along the way with PE6 associating this VPN label with the connection to CE8.

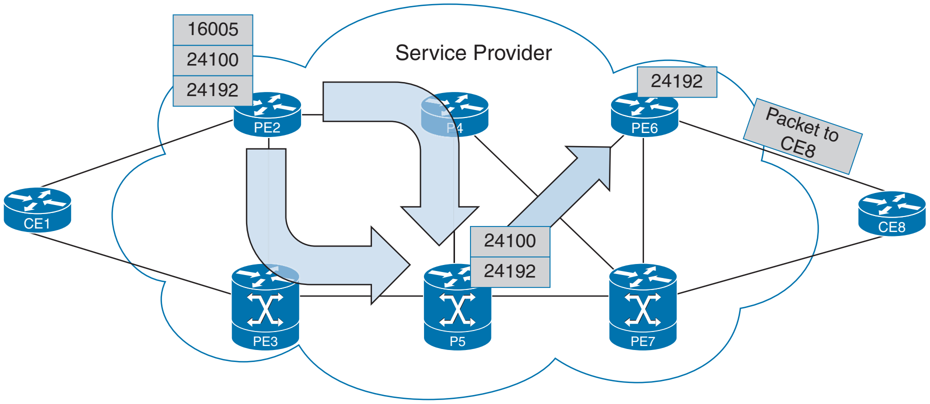


Figure 15-6 *Label Stack Example*

Segment Routing Control Plane

The control plane in Segment Routing (SR) plays a crucial role in managing how segment ID information is shared among network devices. Link-state Interior Gateway Protocol mechanisms distribute segment IDs on Segment Routing networks. Both OSPF and IS-IS include protocol extensions to support the distribution of segment IDs. These extensions enable routers to maintain a comprehensive database containing information about all nodes and adjacency segments. Because IGPs are now responsible for distributing segment IDs, and labels in the case of the MPLS data plane, there's no need for a separate label distribution protocol, as mentioned earlier. Our control plane has become far simpler because it is working with only one source of truth—IGP—instead of having to reconcile both IGP and LDP information during failure events. It is important to note that the **Segment Routing control plane** can be applied to both MPLS and IPv6 data planes. In Cisco's documentation, this is referred to as SR-MPLS and SRv6, the former running on MPLS labels and the latter on IPv6 routing. Let's start by examining SR-MPLS and learn the details behind protocols that provide this unified well-organized improvement.

IS-IS Control Plane

The **IS-IS control plane** disseminates Segment Routing information within an autonomous system. Because LDP is not necessary, IS-IS will distribute both the prefixes and labels in the extensions built into IS-IS itself. This allows for seamless deployment of Segment Routing in existing MPLS networks. Rather than modifying the protocol itself, the designers "extended" its use by providing these additional protocol add-ons to carry information not originally intended by protocol designers. Think of a train of railway cars where the locomotive does not know the load being carried inside each car it is pulling. This way, new functionalities can be added to the protocol by adding new TLVs. IS-IS works exactly this way, because it understands how to transport such values for the use of Segment Routing. It uses Type-Length-Value (TLV) triplets along with sub-TVLs to encapsulate various information in its advertisements. It can support both IPv4 and IPv6 control planes and extends its reach to level-1, level-2, and multilevel routing. It is capable of providing MPLS penultimate hop popping (PHP) and explicit-null signaling as well. Several RFCs, including RFC 8667 and RFC 8402, describe the process of how Prefix-SID and Adj-SID are carried in sub-TLVs in great detail. Figure 15-7 shows the format of the Prefix-SID sub-TLV.

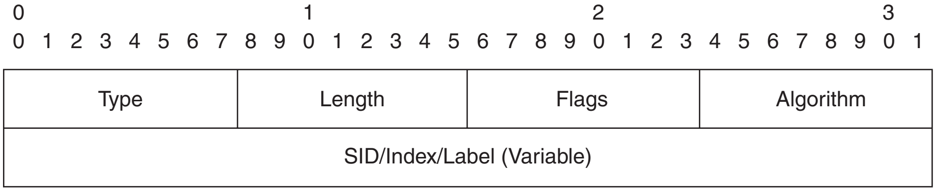


Figure 15-7 *IS-IS Prefix-SID Format*

Table 15-3 shows the most significant TLVs you should be able to recognize on the exam.

Table 15-3 IS-IS TLVs

TLV	Name	Description	Reference
2	IIS Neighbors	Shows all running interfaces to which IS-IS is connected, has a maximum metric of 6 with only 6 out of 8 bits used.	ISO 10589
10	Authentication	The information is used to authenticate IS-IS PDUs.	ISO 10589
22	Extended IS Reachability	Increases the maximum metric to 3 bytes (24 bits), addressing TLV 2 metric limitation.	RFC 5305
134	TE Router ID	MPLS Traffic Engineering router ID.	RFC 5305
135	Extended IP Reachability	Provides a 32-bit metric with an “up/down” bit for the route-leaking of L2 Ð L1.	RFC 5305
149	Segment Identifier/Label Binding	Advertises prefixes to SID/Label mappings. This functionality is called the Segment Routing Mapping Server (SRMS).	RFC 8867
222	MT-ISN	Allows for multiple-topology adjacencies.	RFC 5120
236	IPv6 Reachability	Describes network reachability through the specification of a routing prefix.	RFC 5308
242	IS-IS Router CAPABILITY	Allows a router to announce its capabilities within an IS-IS level or the entire routing domain.	RFC 7981

IS-IS allocates the SRGB along with the Adjacency-SIDs and advertises both in IS-IS for the enabled address-families. IS-IS enables MPLS forwarding for all non-passive interfaces. Example 15-3 shows commands necessary to turn on Segment Routing in IS-IS.

Example 15-3 *Commands to Turn on Segment Routing in IS-IS*

```
IOS XR
RP/0/0/CPU0:PE4# show running-config router isis
router isis CCNP
 set-overload-bit on-startup 300
 is-type level-2-only
 net 49.0001.0000.0000.0004.00
 distribute link-state
 nsf ietf
 log adjacency changes
 lsp-gen-interval maximum-wait 10000 initial-wait 20 secondary-wait 200 level 2
 lsp-refresh-interval 65000
 max-lsp-lifetime 65535
 address-family ipv4 unicast
  metric-style wide
  metric 100 level 2
  microloop avoidance
  mpls traffic-eng level-2-only
  mpls traffic-eng router-id Loopback0
  spf-interval maximum-wait 2000 initial-wait 50 secondary-wait 200
  router-id Loopback0
  segment-routing mpls
 !
 address-family ipv6 unicast
  metric-style wide
  spf-interval maximum-wait 2000 initial-wait 50 secondary-wait 200
 !
 interface Loopback0
  passive
  address-family ipv4 unicast
   prefix-sid index 4
  !
 !
mpls traffic-eng
RP/0/0/CPU0:PE4#
```

OSPFv2 Control Plane

Much like in IS-IS, OSPF does not rely on LDP to transmit prefix label information. It uses protocol extensions to distribute Segment Routing labels in the **OSPFv2 control plane**. OSPF relies on fixed-length link-state advertisements (LSAs) for its fundamental operations. Later, Opaque LSAs were introduced to expand to new protocol capabilities, accommodating features like Segment Routing and Traffic Engineering. These advertisements are flooded to OSPF neighbors opaquely, implying that even if a transit router lacks comprehension of this information (perhaps due to running older software), it will nonetheless indiscriminately transmit it to its neighboring routers. Multi-area functionality is supported, host loopback prefixes are advertised as IPv4 Prefix Segment IDs (Prefix-SIDs), and Adjacency Segment IDs (Adj-SIDs) are used for adjacencies. MPLS penultimate hop popping (PHP) and explicit-null signaling are also supported.

Note the format of Opaque LSAs in Figure 15-8.

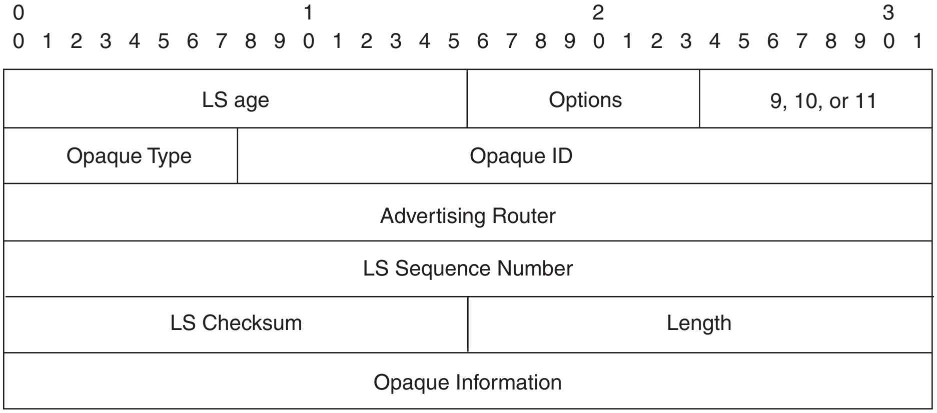


Figure 15-8 *Opaque LSA Format*

Opaque LSA types are identified by the topology flooding scope in Table 15-4. The most known of these (when it comes to Segment Routing) is type 10 LSAs, which distribute Traffic Engineering (TE) link attributes. It is often referred to as the *TE LSA*, yet it has other applications as well.

Table 15-4 OSPF Opaque LSAs

LSA type	LSA Scope	Topology Flooding Scope	Reference
9	Link-local	Local network only	RFC 5250
10	Area-local	Only within an area	RFC 5250
11	Autonomous System	Domain-wide, same as AS-External type-5 LSAs	RFC 5250

Similar to IS-IS, OSPF will allocate and advertise the SRGB to its neighbors. It activates MPLS forwarding on all OSPF interfaces, excluding loopback interfaces, and assigns Adjacency-SIDs to these interfaces. Example 15-4 shows commands necessary to turn on OSPF for Segment Routing.

Example 15-4 *Commands to Turn on OSPF for Segment Routing*

```
IOS XR
RP/0/0/CPU0:PE4# show running-config router ospf
router ospf CCNP
 nsr
 distribute link-state
 log adjacency changes detail
 router-id 10.1.100.10
 segment-routing mpls
 segment-routing forwarding mpls
 fast-reroute per-prefix
 fast-reroute per-prefix ti-lfa enable
 affinity-map
  RED bit-position 0
 !
 nsf ietf
! Output omitted for brevity
 address-family ipv4 unicast
 area 0
  mpls traffic-eng
  segment-routing mpls
  interface Loopback0
   passive enable
   prefix-sid index 10
!
  interface HundredGigE0/0/0/0
   bfd minimum-interval 20
   bfd fast-detect
   bfd multiplier 3
   cost 200
   network point-to-point
  !
  interface HundredGigE0/0/0/1
   bfd minimum-interval 20
   bfd fast-detect
   bfd multiplier 3
   cost 200
   network point-to-point
  !
 !
mpls traffic-eng
RP/0/0/CPU0:PE4#
```

What about OSPFv3? While OSPFv3 has the potential to accommodate Segment Routing for IPv6 and utilize a native IPv6 data plane, specific extensions outlined in an IETF draft are required for implementation. It's worth noting that, at press time, these extensions have not been integrated into Cisco IOS XR and IOS XE.

BGP Control Plane

BGP also has the capability to function as the control plane for Segment Routing (SR), enabling prefix distribution throughout the network. In the context of Segment Routing, the **BGP control plane** distributes segment routing information between routers, enabling them to make forwarding decisions based on predefined segments. While seen less frequently than IS-IS and OSPF, BGP has been effectively used in practice in multiple large-scale web data centers. Such data centers can support over 100,000 services, profoundly influencing and challenging the scalability and operational efficiency of the underlying network architectures. To meet the demands of high-intensity east-west traffic found in these compute clusters, operators frequently opt for variations of Clos or Fat-tree topologies. In these massive data center networks, symmetrical topologies with numerous parallel paths connecting two server-attachment points are common. It is in this context that BGP excels and arguably surpasses the IGP approach. The assertion that "BGP is a better IGP" challenges traditional viewpoints and has sparked conversations.

What would make BGP an attractive choice? Remember that these massive data centers seek maximum bandwidth to be transferred across the midpoint of the system. Such network structures are designed to be both highly scalable, cost-effective, and are constructed from affordable, low-end access-level switches. To maintain this level of scale, the designs call for a single protocol with simple behavior and wide vendor support. With the above in mind, when it comes to simplicity, BGP certainly has its advantages because it has less of a state machine and fewer data structures. This may not appear intuitive at first glance, but it does not take long to realize that the BGP RIB structure is simpler than those of Link-State Databases (LSDBs). There is a very clear picture of "which routing information is sent where." There is a RIB-In and RIB-Out, a far easier construct for tracing exact routing paths than following link-state topology constraints with areas and levels. When it comes to operational troubleshooting, this is definitely a strength. Also, event propagation is more constrained in BGP because link failures have limited propagation scope. We can argue that BGP has more stability due to the reduced "event-flooding" domains. When it comes to traffic steering, BGP allows for per-hop Traffic Engineering that can be used for unequal cost Anycast load-balancing. In addition, BGP is widely supported by practically all vendors, so from the perspective of interoperability, BGP beats IGPs. We have been conditioned to perceive BGP as slow and suitable primarily for inter-domain routing, but it has no issues with demonstrating its adaptability and effectiveness in modern topologies. Therefore, it is advisable to approach BGP with an open mind, recognizing its potential to perform as well as, or even better than, traditional IGP alternatives in contemporary implementations.

BGP will advertise a BGP Prefix-SID associated with a prefix via BGP Labeled Unicast (BGP-LU) IPv4/IPv6 Labeled Unicast address-families. BGP Prefix-SID is a global SID, and the instruction forwards the packet over the ECMP-aware BGP best path to the associated prefix. RFC 8277 specifies that Label-Index TLV must be present in the BGP Prefix-SID attribute attached to IPv4/IPv6 Labeled Unicast prefixes. This 32-bit value represents the index value in the SRGB space and has the format illustrated in Figure 15-9.

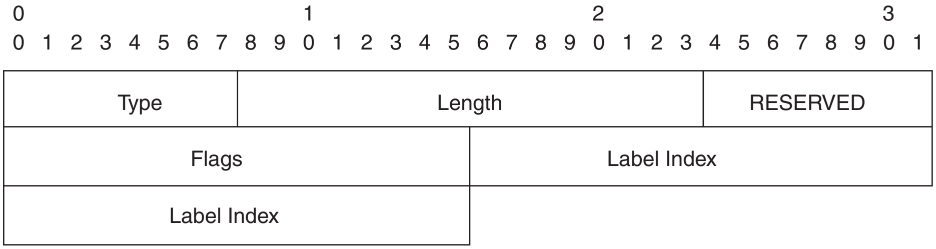


Figure 15-9 *BGP Prefix SID Advertised Format*

The Prefix-SID for a locally originated BGP route can be set with a route-policy. Example 15-5 shows how to attach a label-index with **network** and **redistribute** commands.

Example 15-5 *Attaching a Label-Index via a Route-Policy*

```
IOS XR configuration with network command

route-policy SIDs($SID)
  set label-index $SID
end-policy
!
router bgp 100
 address-family ipv4 unicast
  network 10.1.100/4/32 route-policy SID(1)
  allocate-label all
```

```
IOS XR configuration with redistribute command

route-policy SIDs
 if destination in (10.1.100.4/32) then
  set label-index 1
 endif end-policy
!
router bgp 100
 address-family ipv4 unicast
  redistribute connected route-policy SIDs
  allocate-label all
```

One last thing regarding having BGP for a Segment Routing control plane. Remember the Anycast load-balancing I mentioned earlier in this section? Anycast allows different nodes to advertise the same BGP prefix. It is an application of Prefix SIDs to achieve anycast operations. Look at Figure 15-10, where I again moved some links around to represent a data center's spine-and-leaf architectures, with spines located at the top. PE2 and P4, while advertising their individual BGP Prefix-SIDs (16002 and 16004, respectively), have been made members of the same unicast set. Both of them advertise anycast prefix 10.1.100.24/32 with BGP-Anycast SID 20001. PE3 wants to send traffic to PE7 but would like to exclude spine PE6. BGP-Anycast SID 20001 will load-balance the traffic to any member of the Anycast set and then forward it to PE7.

Additionally, due to BGP Prefix-SID global label usage, BGP-LU local labels are going to be the same across all of the network's ASBRs. As a result, these Anycast loopbacks can be used as the next-hop for BGP-LU prefixes. That is pretty good resiliency! Nothing to scoff at, for sure.

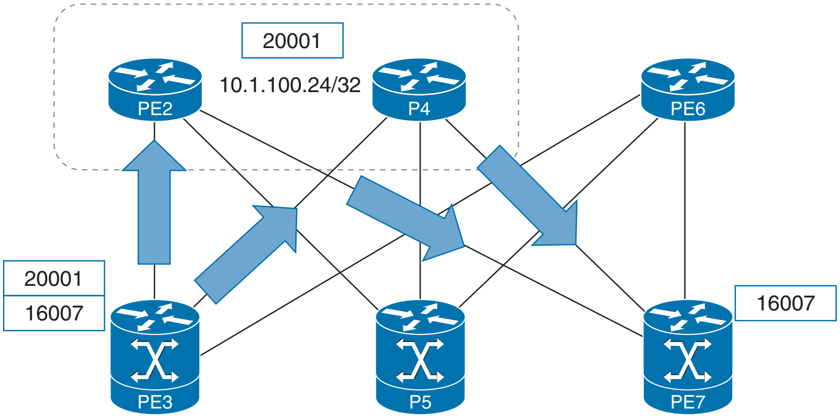


Figure 15-10 *BGP-SR Anycast Load-Balancing*

SRv6 Control Plane

The **SRv6 control plane** manages the signaling, routing, and forwarding information for Segment Routing over IPv6 (SRv6) networks. It serves as the Segment Routing architecture tailored for the IPv6 data plane and extends to the value of IPv6, influencing future IP infrastructure deployments, whether in data centers, large-scale aggregation, or backbone networks. SRv6 functions as an extension of the Segment Routing architecture specifically designed for IPv6 networks. It introduces a source-routing mechanism by encoding instructions within the IPv6 packet header.

The use of IPv6 addresses to identify objects, content, or functions applied to objects opens up significant possibilities, particularly in the realm of chaining microservices within distributed architectures or optimizing content networking. Notably, stable networks, particularly in the Asia-Pacific region, have embraced SRv6, boasting tens of thousands of nodes on a single network as of the time of writing this book.

Fundamentally, SRv6 encodes topological and services paths into the packet header. The SRv6 domain does not hold any per-flow state for Traffic Engineering or network function virtualization (NFV). Sub-50ms path protection is delivered with TI-LFA. It natively delivers all services in the packet header, without any shims or overlays. IPv4's limitations have forced the industry to create extra tools to deal with its challenges. When IPv4 lacked sufficient address space, NAT was created to hide and conserve addresses. For engineered load-balancing, we have had to invent MPLS Entropy Label and VxLAN UDP. For separating discrete networks, MPLS VPNs along VxLAN were created. Since Traffic Engineering functions were missing in IPv4, RSVP-TE and SR-TE MPLS appeared. Network Service Header (NSH) overcame IPv4 service chaining limitations. All of the above is done natively in IPv6 and why so many service providers are turning to this technology.

SRv6 (Segment Routing over IPv6) Header

At the heart of SRv6 is the IPv6 Segment Routing Header (SRH). Figure 15-11 shows the IPv6 SRH replicated from RFC 8754. This header is added to IPv6 packets to implement Segment Routing on the IPv6 forwarding plane. SRH specifies an IPv6 explicit path, listing one or more intermediate nodes the packet should visit on the way to its final destination. The Segment Left field provides the number of transit nodes before traffic reaches its destination. Then, the Segment List fields indicate the sequence of nodes in 128-bit IPv6 addresses to be visited from bottom to top. Segment List [n] shows the first node in the path; Segment List [0] shows the last node in the path.

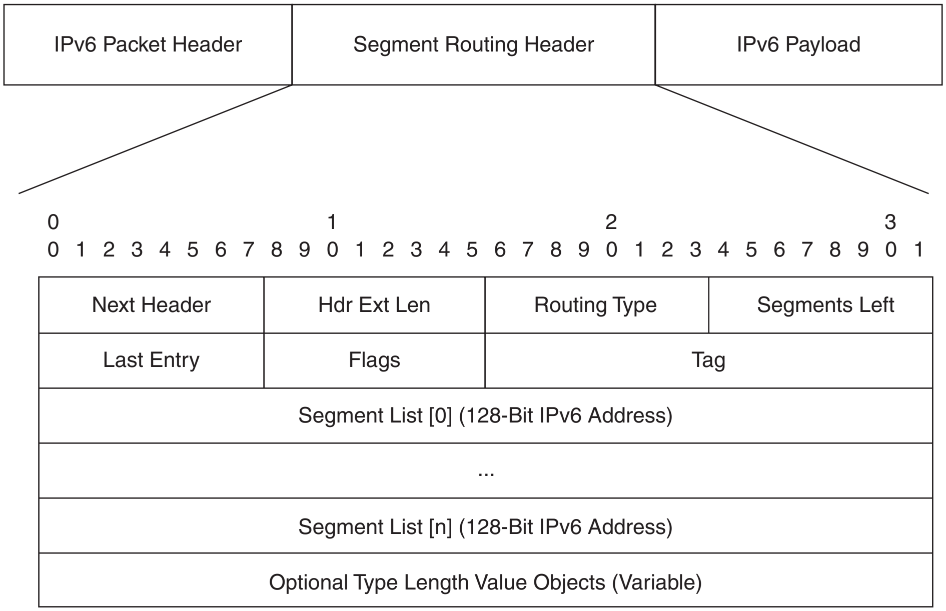


Figure 15-11 *IPv6 Segment Routing Header Format*

In SRv6, each segment is represented by an IPv6 address known as a segment identifier (SID). These SIDs play a crucial role in defining specific paths or instructions for forwarding packets throughout the network. It looks a lot like a 128-bit IPv6 address, but has different semantics because it consists of two parts, with Figure 15-12 providing the visualization:

- **Locator:** Represents an address of a specific SRv6 node performing the function.
- **Function:** Represents any possible network instruction bound to the node that generates the SRv6 SID (network instruction) and is executed locally on that particular node, specified by the locator bits.

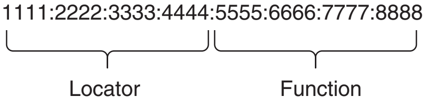


Figure 15-12 *IPv6 Segment Identifier*

You now have the ability to send packets to a node (locator) and then instruct the node to execute an action (function). This is not a subtle difference! In SR-MPLS, IGP with extensions advertised the transport mechanism, and services (L2VPNs, L3VPNs) were signaled *independently* via LDP or MP-BGP. You could change your transport (from MPLS to SR) without affecting the upper protocols that ran on top of it. For the first time in the industry, transport and services instructions are coupled and signaled in the SID. You will see an example of this coming up shortly where an L3VPN is written into the SID.

SRv6 Node Roles

In the context of SRv6 (Segment Routing over IPv6) networks, different nodes play distinct roles in facilitating packet forwarding and processing. These roles include

- **Source Node:** This node has the capability to generate an IPv6 packet incorporating a Segment Routing Header (SRH), essentially forming an SRv6 packet. Alternatively, it serves as an ingress node that can apply an SRH to an existing IPv6 packet.
- **Transit Node:** Found along the SRv6 packet's path, the transit node functions without inspecting the SRH. The destination address of the IPv6 packet does not align with the transit node, and its role is primarily to forward the packet.
- **Endpoint Node:** Located within the SRv6 domain, this node acts as the termination point for the SRv6 segment. The destination address of the IPv6 packet containing an SRH corresponds to the endpoint node. The endpoint node executes the specific function associated with the SID bound to the segment.

SRv6 Micro-Segment SID (uSID)

Often referred to as *Micro-SID* or *Compressed SID*, the uSID feature is an extension of the SRv6 architecture. In SRv6, the micro segment identifier, or uSID, is a specialized form of Segment Routing where packets are marked with a compact identifier for precise forwarding. Unlike traditional SRv6, which might use longer segment identifiers for various purposes, uSID is specifically designed for efficient and granular traffic steering. It provides a more streamlined approach to segment routing, particularly useful for scenarios requiring fine-grained control and scalability enhancements.

Using the established SRv6 Network Programming framework, it can encode up to six SRv6 Micro-SID (uSID) instructions within a singular 128-bit SID address, termed a *uSID Carrier*. Moreover, this extension seamlessly integrates with the existing SRv6 data plane and control plane, requiring no modifications. Notably, it ensures minimal MTU overhead. For instance, when incorporating six uSIDs within a uSID carrier, it yields 18 source-routing waypoints with just 40 bytes of overhead in the Segment Routing Header. Look at Figure 15-13, which illustrates the usage of uSID. Pay attention to how the highlighted uSIDs correspond to router numbering/naming.

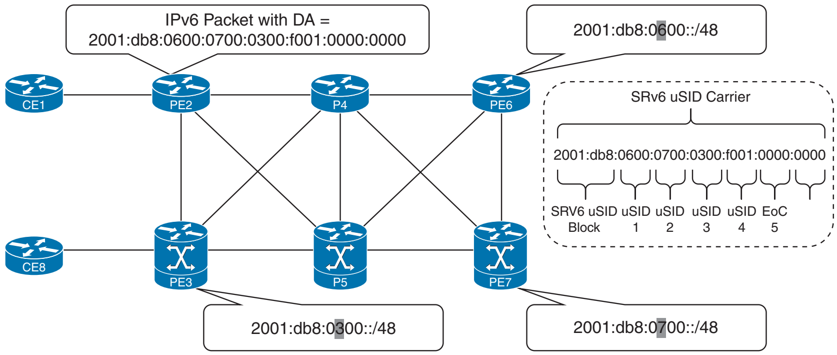


Figure 15-13 *uSID in Action*

The customer at CE1 is using the VPNv4 SP service to connect to a remote site CE8. Router PE2 sends traffic to VPNv4 CE8 to router PE3 via a traffic-engineered path visiting routers

PE6 and PE7 using a single (!) SRv6 SID (note that without uSID, a sequential Segment List would have to be specified). Let's unpack this:

1. PE2, PE6, PE7, and PE3 are SRv6 capable and are configured with 32-bit SRv6 block 2001:db8.
2. P4 and P5 run classic IPv6 forwarding and do not change the Destination Address.
3. PE6, PE7, and PE3 advertise their corresponding 2001:db8:0600::/48, 2001:db8:0700::/48, and 2001:db8:0300::/48 routes.
4. PE2 receives an IPv4 packet from CE1, encapsulates it, and sends an IPv6 packet with the destination address 2001:db8:0600:0700:0300:f001:0000:0000. This is an SRv6 uSID Carrier that contains a sequence of micro-SIDs (instructions 0600, 0700, 0300, f001, and 0000).
5. The 0600, 0700, and 0300 uSIDs are used to construct a traffic engineering path to PE3 with two stops along the way—PE6 and PE7. uSID f001 is a BGP-signaled instruction sent by PE3 indicating the VPNv4 service. uSID 0000 indicates the end of instructions.
6. What happens at P4? P4, running only classic IPv6, forwards the packet along the shortest path to PE6.
7. PE6 receives the packet, pops its own uSID 0600, and advances the micro-program by looking up the shortest path to the next Destination Address (DA) 2001:db8:0700::/48. Now the DA is 2001:db8:0700:0300:f001:0000:0000:0000. This behavior is called *shift and forward*.
8. PE7 receives the packet, pops its own uSID 0700, and advances the micro-program by looking up the shortest path to the next Destination Address (DA) 2001:db8:0300::/48. Now the DA is 2001:db8:0300:f001:0000:0000:0000:0000. Shift and forward again.
9. P5 forwards the packet to PE3, just like P4 did.
10. PE2 receives the packet and executes the VPNv4 function based on this own instruction f001. It decapsulates the IPv6 packet, performs IPv4 table lookup, and forwards the IPv4 packet to CE8.

SRv6/MPLS L3 Service Interworking Gateway

The SRv6/MPLS L3 Service Interworking Gateway facilitates the seamless extension of L3 services between MPLS and SRv6 domains, ensuring continuity in service delivery across both control and data planes. This feature enables interoperability between SRv6 L3VPN and existing MPLS L3VPN domains, offering a pathway for transitioning from MPLS to SRv6 L3VPN.

At the gateway node, the SRv6/MPLS L3 Service Interworking Gateway performs both transport and service termination tasks. It generates SRv6 VPN SIDs and MPLS VPN labels for all prefixes within the configured VRF for re-origination, as illustrated in Figure 15-14. The gateway supports traffic forwarding from the MPLS domain to the SRv6 domain by removing the MPLS VPN label, performing a destination prefix lookup, and applying the appropriate SRv6 encapsulation. Conversely, for traffic from the SRv6 domain to the MPLS domain, the gateway removes the outer IPv6 header, performs a destination prefix lookup, and applies the VPN and next-hop MPLS labels.

PE3 is the interworking gateway that has one leg in the SR-MPLS domain and the other in the SRv6 domain. It performs a translation service by popping the MPLS VPN label and looking up the destination prefix in the SRv6 domain. It encapsulates the payload in the outer IPv6 header with P4's destination address. In the opposite direction, PE3 removes the outer IPv6 header, looks up the destination prefix, and pushes MPLS label 16002 for the BGP next-hop of PE2.

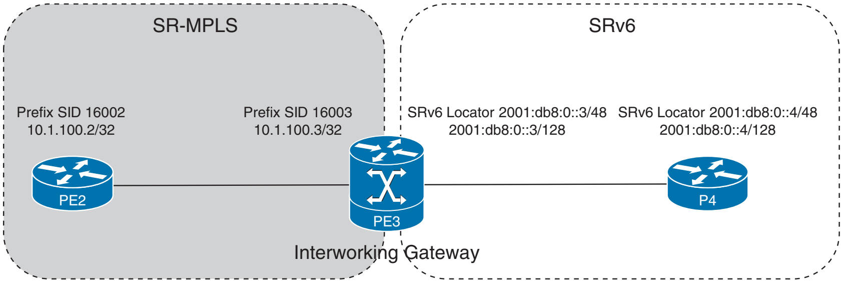


Figure 15-14 *SR-MPLS SRv6 Interworking Gateway*

Co-existence with LDP

It would be nice to never worry about LDP and RSVP, but the reality is that many of today's engineers will have to touch these older MPLS networks. A Segment Routing control plane can co-exist with the label-switched paths (LSPs) constructed with LDP or RSVP. The MPLS architecture allows for the simultaneous use of multiple label distribution protocols, including LDP, RSVP-TE, and others. The SR control plane can coexist alongside these protocols without any interaction. In Figure 15-15, we have removed some links in our network and have thus completely flattened it. This network runs a mix of both Segment Routing (SR) and Label Distribution Protocol (LDP). It is possible to establish an end-to-end seamless Multiprotocol Label Switching (MPLS) LSP, which will ensure interoperability between these two domains. To accomplish this, one or more nodes function as Segment Routing Mapping Servers (SRMS). These SRMS entities take on the responsibility of advertising SID mappings on behalf of nodes that are not SR-capable. This mechanism enables SR-capable nodes to learn about the SIDs assigned to non-SR-capable nodes without the need for explicit individual node configurations. Let's unpack this.

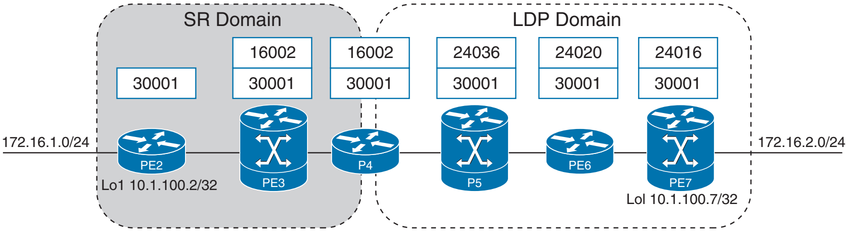


Figure 15-15 *SR and LDP Domain Interoperability*

Notice that this network runs both SR and LDP, which can be typical during network transitions and upgrades. PE2 and PE7 are exchanging BGP VPNv4 routes. PE2, PE3, and P4 are SR-capable. PE4, P5, PE6, and PE7 use LDP. How do these two domains talk to each other end to end? First, let's start from the LDPÐSR direction, which is quite easy because SR-capable routers will automatically translate between LDP- and SR-based labels:

1. PE7 learns a service route (L3VPN route, for example) for customer prefix 172.16.1.0/24 with a service/VPN label of 30001.
2. PE7's BGP next-hop for this service label is associated with PE2's lo1 10.1.100.2/32.

3. PE7 finds LDP label binding 24016 from its neighbor PE6 for PE2's Forwarding Equivalence Class (FEC) of 10.10.100.2/32 and forwards the packet to PE6.
4. PE6 finds LDP label binding 24020 from its neighbor PE5 for PE2's FEC of 10.10.100.2/32, swaps 24016 for 24020, and forwards the packet to PE5.
5. PE5 finds LDP label binding 24036 from its neighbor PE4 for PE2's FEC of 10.10.100.2/32, swaps 24020 for 24046, and forwards the packet to P4.
6. P4 lacks an LDP binding originating from its next-hop PE3 for the FEC associated with PE1. What it does carry, though, is an SR node segment pointing to an IGP route leading to PE2. P4 engages in label merging, wherein it replaces its local LDP label (24036) for FEC PE2 with the corresponding SR node segment label, which is 16002.
7. PE3 pops label 16002 (assuming penultimate hop popping function is used) and forwards the packet to PE2.
8. PE2 receives the packet, looks up its service label of 30001, and drops the packet into the appropriate customer VRF.

We now have an end-to-end LDPÐSR path. Simple. What about in the opposite direction? This is where we will encounter a problem going from SRÐLDP. Can you take a moment to think what the problem would be by looking at Figure 15-14 before you examine Figure 15-15? PE2 needs to send traffic to 172.16.2.0/24 with service label 40001 that it received with the BGP next-hop of 10.1.100.7/32. Since PE2 only speaks SR, when it looks up the node segment for 10.1.100.7/32, what will it find in its label database? Nothing. Why? Because such label mapping does not exist on the network, since the operator has never configured it; therefore, no router advertises or receives this label mapping. There must be something to associate PE7's loopback with SR label mapping. The better answer here is Segment Routing Mapping Server (SRMS). All analogies finally break down, but it is possible to think of SRMS as a sort of route reflector for SR labels. Just as in BGP, we can centrally instruct all routers in our SR domain. Look at Figure 15-16.

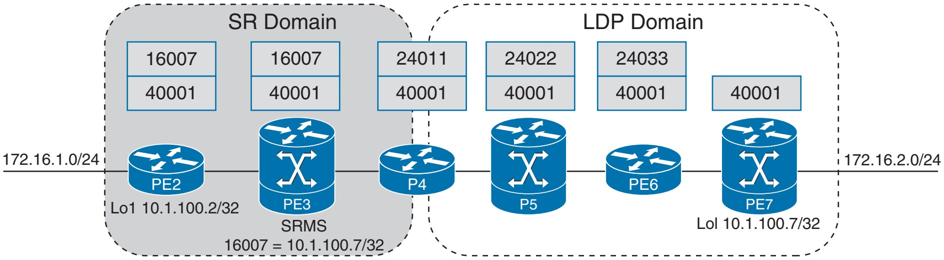


Figure 15-16 *SR and LDP Domain Interoperability with SRMS*

Walking back in the opposite direction looks like this:

1. PE3 is chosen as a Segment Routing Mapping Server (SRMS). In practice, it is recommended to have a redundant SRMS.
2. As PE7 lacks Segment Routing (SR) capability, you must create a mapping policy on the SRMS, which associates label 16007 with PE7's lo1 10.1.100.7/32.
3. Now, PE2 learns a service route (L3VPN route via BGP) for customer prefix 172.16.1.0/24 with a service/VPN label of 40001 with the BGP next-hop of 10.1.100.7/32.

4. PE2 finds an SR label binding 16007 it has received from the SRMS PE3 for PE7's FEC of 10.10.100.7/32 and forwards the packet to PE3 as the IGP next-hop.
5. PE3 finds an SR label binding 16007 pointing to its neighbor P4 as the IGP next-hop, swaps 16007 for 16007, and forwards to P4.
6. P4 does not have an SR label for PE7's IGP route, but it holds LDP label 24011 for this FEC. It swaps 16007 for 24011 (remember the process is called *label merge*) and forwards to P5.
7. P5 swaps 24011 for 24022 and forwards to PE6.
8. PE6 pops the label (due to PHP in this setup) and forwards to PE7.
9. PE7 receives the packet, looks up its service label of 40001, and drops the packet into the appropriate customer VRF.

What should you remember here? Segment Routing Mapping Server labels are only necessary in the SRÐLDP direction. SR and LDP labels come from separate label database ranges (16000–23999 for SR and 24000+ for LDP), so unless the operator has deliberately violated this guidance, there is not a chance the network will be in a state of confusion, since the SR and LDP labels do not overlap each other. The network must maintain continuous SR connectivity in the SR domain. The network must also maintain continuous LDP connectivity in the LDP domain. If you understood the packet walkthrough, these points should be clear.

One last thing to know for completeness. By default, Cisco routers prefer LDP as the label imposition mechanism when the MPLS features are turned on. The way to enable SR for label imposition is shown in Example 15-6.

Example 15-6 *Segment Routing Label Imposition Preferred*

```
IS-IS

router isis 100
 address-family ipv4|6 unicast
  segment-routing mpls sr-prefer
```

```
OSPF

router ospf 100
 segment-routing mpls
 segment-routing sr-prefer
```

Segment Routing Traffic Engineering

RSVP-TE, despite its powerful Traffic Engineering capabilities, poses challenges in practical deployments due to its complexity. Managing backup tunnels, intricate configurations at scale, the absence of seamless inter-domain intelligence, and the complexities of steering traffic through methods like PBR or autoroute have resulted in various issues and limited widespread adoption. Simplifying these aspects is crucial for enhancing the usability and deployment scope of Traffic-Engineered networks. Enter Segment Routing policies. They are simple, automated, scalable, and carry support for a wide variety of functionalities including multidomain intelligence, which is provided by Path Computation Element (PCE) and Binding-SID (BSID)—more on this later.

Segment Routing Policies

In Segment Routing, there are no tunnels (the closest possible thing is Circuit-Style Segment Routing, which has policies to put traffic on the same A–Z path, akin to bidirectional co-routed LSPs—this is outside of the current exam's scope). Instead, Segment Routing introduces the concept of Segment Routing Policies. These are typically deployed at ingress routers at the edge of the network and can force the packet to follow any desired path.

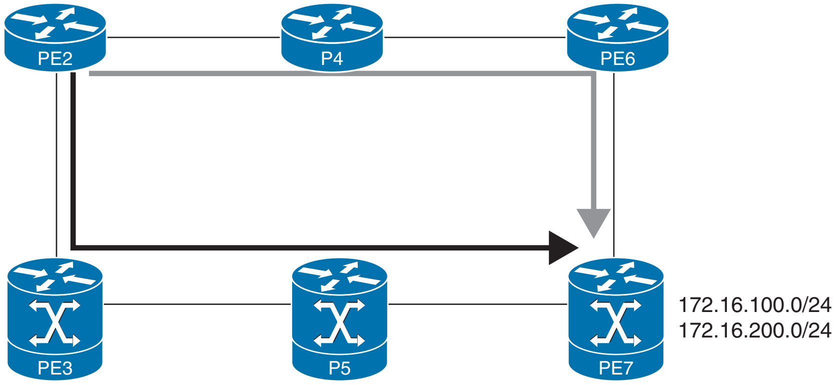
Key Topic

An SR Policy is fundamentally a sequence of segments. In its most basic structure, it is a sequence of IP waypoints presented in either SR-MPLS or SRv6 format (SID list), with the initial entry as the first destination to be visited. An SR Policy is uniquely identified by these attributes:

- **Headend:** An ingress router where the policy is implemented.
- **Tailend:** An egress router where the policy ends.
- **Color:** A numeric value that uniquely identifies multiple SR Traffic Engineering policies between the same pair of routers.

Figure 15-17 illustrates this best. PE2 needs to send traffic for prefixes 172.16.100.0/24 and 172.16.200.0/24 to the same PE7 router, since it is the egress point connecting these two networks. However, traffic destined for 172.16.100.0/24 must follow the top low-delay path due to latency requirements, and traffic for 172.16.200/24 must take the bottom low-cost path because the customer is not paying for the premium service. Try doing this with IGP alone! SR policies, on the other hand, easily differentiate traffic between the same pair of routers by steering them into differently colored policies (different numeric values) that properly groom traffic onto the desired paths.

SR Policy "Gray" for Low-Delay Paths:
1) Headend = PE2
2) Tailend = PE7
3) Color = Gray (Numeric Value 100)



SR Policy "Black" for Low-Cost Paths:
1) Headend = PE2
2) Tailend = PE7
3) Color = Black (Numeric Value 200)

Figure 15-17 *Segment Routing Policy Places Traffic on Diverse Paths*

SR Policies and Candidate Paths

An SR Policy consists of one or more **Candidate Paths**. Each Candidate Path has a single SID-list or a set of weighted SID-list. Things to consider regarding a Candidate Path (the order is not important here):

1. Can be explicitly defined. The operator will provide the exact sequence of SIDs to be visited along the way to the destination.
2. Can be dynamically defined. The operator will provide optimization objectives (select only encrypted links) and constraints, a set of rules to follow (exclude links certain attributes, such as not meeting minimum delay).
3. Has a preference value (numeric, higher is preferred).
4. Is associated with a single Binding-SID (BSID, more on this later).
5. Can be supplied to the headend via
    a. CLI
    b. NETCONF
    c. PCEP (Path Computation Element Protocol)
    d. BGP
6. An SR Policy will select a single best Candidate Path and program it via BSID into the router's RIB/FIB forwarding table.

Key Topic

Binding-SID (BSID)

Binding Segment Identifier (BSID) is a SID value that is an opaque representation of a Segment Routing Policy. BSID shows a chosen path to upstream routers. It provides isolation and decoupling between distinct source-routed domains while increasing overall network scalability. Do not forget that SR Policies use BSID to program a router's forwarding table (just mentioned in point 6).

Note how in Figure 15-18 different routing/SR domains are involved. The list of SIDs to steer traffic onto the imagined low-delay path between DC1 and DC5 (DC2, PE2, P4, P4-ADJ-SID, PE7, DC5) can be long. A single BSID can represent the entire Segment Routing Policy sending it through the WAN Core domain, requiring only three SIDs (DC Primary, WAN Core SR Policy, DC Secondary). This reduces the number of segments imposed by the source.

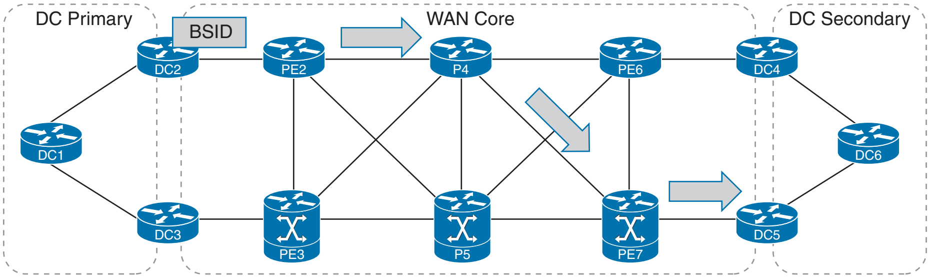


Figure 15-18 *Multidomain Use of Binding-SID*

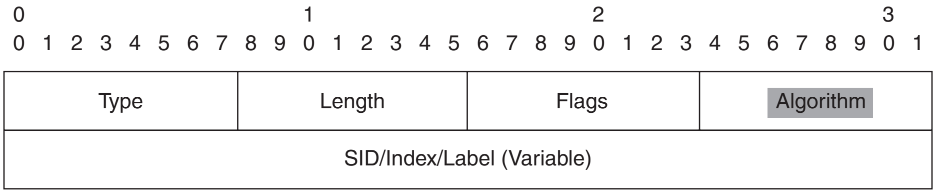
Additionally, this approach keeps one domain unaffected by routing changes in another domain, since BSID does not change during these events. Domain internal operations can be thus hidden (opaque) from each other, which can be beneficial to service providers who do not want to disclose the details of how they provide services to their customers.

Flex-Algo

Flex-Algo is the best way to do traffic engineering today. Flex-Algo, short for Flexible Algorithm, enhances Segment Routing Traffic Engineering (SRTE) by introducing additional segments with distinct properties compared to the Interior Gateway Protocol Prefix segments. It expands the SRTE capabilities by including customizable, user-defined segments in the toolbox. It can also use Segment Routing on-demand next hop (ODN) and Automated Steering to create traffic-engineered paths based on user intentions; these are outside of the scope of this book.

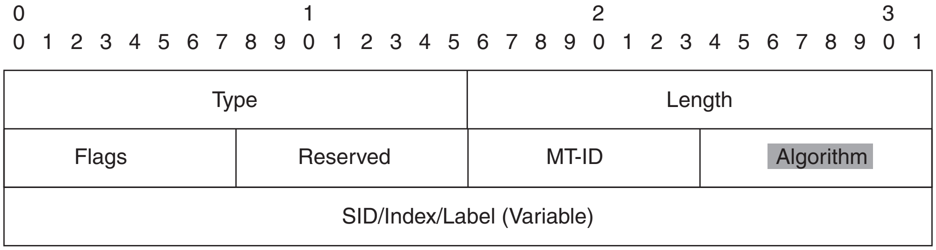
IETF has standardized algorithms 0 through 127. Routers run the default algorithm 0 as the IGP shortest path derived from the IGP metric. Additional algorithms 128 through 255 can be customized by network operators. They are known as *SR IGP Flexible Algorithms*, or Flex-Algo as the shorter version. It is called *flexible* because you can decide which metric you want to use in your intent.

In our earlier discussion of Prefix-SID in this chapter, we focused solely on explaining the default aspect of Prefix-SID behavior, specifically the one linked to algorithm 0. When you read (you really should) RFC 8867 and RFC 8665, you will notice that both IS-IS and OSPF include Prefix-SID sub-TLV algorithm field in the formats illustrated in Figure 15-19 and Figure 15-20.



Type	Length	Flags	Algorithm
SID/Index/Label (Variable)			

Figure 15-19 *Algorithm Field in Prefix-SID Sub-TLV for IS-IS*



Type		Length	
Flags	Reserved	MT-ID	Algorithm
SID/Index/Label (Variable)			

Figure 15-20 *Algorithm Field in Prefix-SID Sub-TLV for OSPF*

This means that the operator can change the default algorithm 0 (IGP shortest-path) behavior on routers that are assigned to use a different algorithm (128–255) as different constraints (logical rules) will be imposed on the part of the network that participates in this algorithm. Let's look at Figure 15-21, where we break from the familiar-to-us topology.

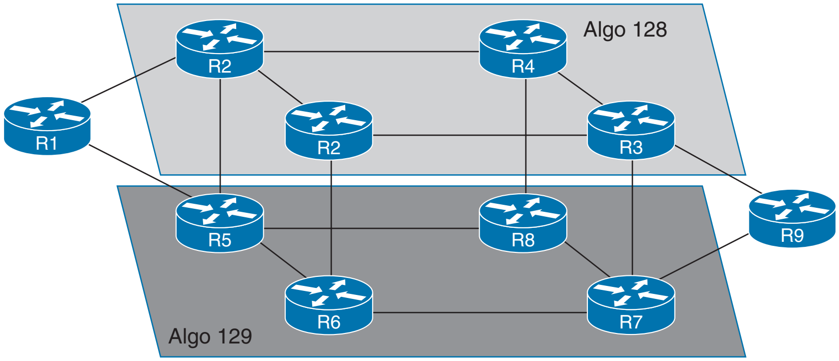


Figure 15-21 *Flex-Algo Network Slicing*

Suppose the operator wanted to impose different type behaviors on this network. Instead of using the default algo 0 (IGP shortest-path), the operator can define other algorithms that can minimize metrics other than shortest path, such as delay, for example. The operator can also combine this with rules to exclude links with certain properties (link-affinity, SRLG, encryption, etc.). Here is one example of what can be done:

1. The operator can define Flex-Algo 128 to prioritize IGP metric and avoid link-affinity "dark-gray" on the bottom.
2. The operator can define Flex-Algo 129 to delay metric and avoid link-affinity "light-gray" on the top.
3. Routers R1 and R9 would be added to participate in algo 0, 128, and 129.
4. Routers R1, R2, R3, and R4 would be added to participate in algo 0, 128.
5. Routers R5, R6, R7, and R8 would be added to participate in algo 0, 129.

Consider how powerful the network has become. The operator has bisected the network into two distinct profiles. The top part has routes based on the shortest path according to the IGP. The bottom half will route delay-sensitive traffic through the part of the network that uses dynamic link-delay measurement, which will be advertised by the IGP. Someone reroutes your optical underlay path? Does not matter. Someone moves a circuit without bothering to notify your department? Does not matter. The network will recalculate the best path according to the intent you had in mind. If you see the beauty of this approach, you will understand that the limitation of what can be done on a network at scale exists only in the minds of its architects. Low-cost delay-optimized paths (my personal SP operator nirvana, because this is where I make money) have now become a reality because we can now finally differentiate services on our infrastructure. It is my opinion that while SR policies are effective for carving dynamic paths on the network, the simplicity and flexibility of Flex-Algo allow the operator to easily slice the network into multiple planes that can be used to carry encrypted, application-dependent, dynamic delay-based, low-cost, and other intent-based traffic. You can even drain all network traffic to the bottom dark-gray plane of the network, run upgrades on the top light-gray plane of the network, and repeat the process again in the other direction—accomplishing zero downtime for your customers.

That is the power of Flex-Algo—operational simplicity and scale. You can finally manage massive networks with a simple picture in mind, rather than constructing hundreds of SR Policies for individual applications. Flex-Algo is applicable to SR-MPLS and SRv6. In the case of SR-MPLS, you will get an extra label for the router's loopback. In SRv6, you get a different locator (recall our earlier discussion on this topic). SP operators are really beginning to heavily use this approach with SRv6 (IPv6).

Key Topic

Something you need to be aware of that is directly called out in the exam blueprint is *encapsulation*. SR-MPLS uses SRTE policies (on-demand or manual) to steer traffic into Flex-Algo. If you want traffic to follow a particular path, you specify a list of SIDs. When you specify the SID associated with Flex-Algo, the traffic takes that specific Flex-Algo plane. In contrast, SRv6 does not use policies. In SRv6, the ingress PE will directly encapsulate traffic based on the Service SID advertised in BGP. That Service SID (remember locator + function?) is a combination for the Algo locator and decapsulation function. Transport and Service become blended and are encoded into the transport intent (Algo locator). Transport intent and Service function are *encapsulated* into the same instruction. SRv6 is a much simpler approach to driving traffic intent.

TI-LFA

To date, TI-LFA is the number one reason why network operators deploy SR: they want an automated way to compute a backup path by IGP. No need to do MPLS-TE (traffic engineering tunnels) for fast reroute (FRR). **Topology-independent loop-free alternate (TI-LFA)** provides a simple, automatic, optimal, and topology-independent sub-50ms per-prefix protection to the network. It can protect Segment Routing, LDP, and IP traffic without relying on the construction of backup tunnels of any sort, as is the case in RSVP-TE. Whether IS-IS or OSPF is used, these protocols precompute a backup path for each active path per IP prefix destination. They run an SPF algorithm for the primary path and then automatically run the SPF again, *excluding* the primary path—deriving the backup path. IGP pre-installs this path in the data plane and immediately uses it once the active destination path is impacted. Be careful with analogies, because they all finally breakdown, but it can be helpful to think of how an EIGRP-feasible successor works. The router already knows what the post-convergence path will look like even before the failure occurs.

Figure 15-22 shows a fundamental TI-LFA operation from router PE2's perspective; once the protected PE4–P2 link fails, traffic is rerouted over the post-convergence path, which is known and preprogrammed before the link failure occurs. The recommendation is to enable this functionality on all routers in your Segment Routing domain. This approach creates automatic backup paths throughout the network without the burden of manually provisioning backup tunnel paths.

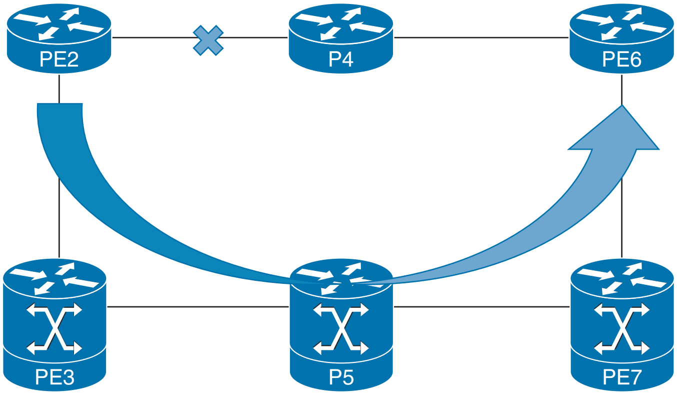


Figure 15-22 *TI-LFA Operation*

Key Topic

Terms from Remote LFA Technology

RFC 7490 describes the following architectural reference areas to understand repair tunnel endpoints for link protection. While there is no concept of tunnels in Segment Routing (they have been replaced by policies), the same reference areas apply and are important for this exam.

P-Space

In Figure 15-23, we return to our topology and remove some of the internal links to create a ring topology to better understand these reference areas.

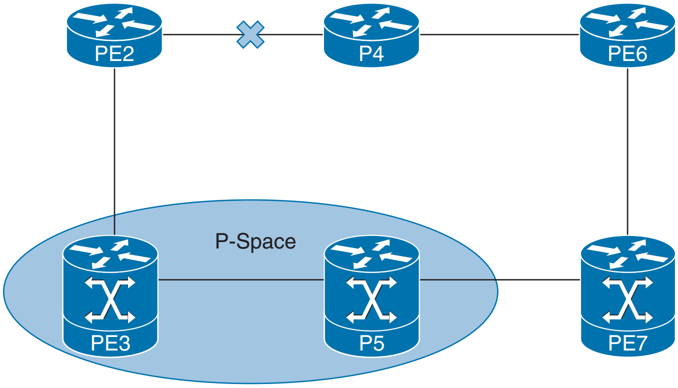


Figure 15-23 *P-Space Reference Area*

Reference areas are always seen from a perspective of a certain router with respect to a particular failed link. The way to look at reference areas depends on which router we're considering and the specific link it is protecting. In this example, router PE2 would like to protect the link between itself and router P4. The protected space (**P-Space**) of a router concerning a protected link refers to routers that PE2 can reach through the shortest paths without having to use the protected link. Which routers would those be? All link costs being equal here, only PE3 and P5 will be in P-Space. What about PE7? Not quite, as it is possible, due to ECMP, that PE2 can send a packet to PE7 through the top of the diagram through the protected link, thus disqualifying from being the shortest path. What would be the point of using a link that can potentially fail? Expressed in cost terms, P-Space contains a set of routers found on a shorter path than the path cost going through the protected link. In the case of PE7, it is equal and not shortest—thus, not a part of P-Space.

Q-Space

Q-space refers to a set of backup paths or alternate next hops that are precomputed for use during a failure. Figure 15-24 shows the other side of the protected PE2–P4 link from router P4's perspective, and the same rules apply again. When following the same rules, the set of routers reachable from P4 via the shortest path without possibly going through the protected link only include PE6 and PE7.

PQ Node

Viable repair tunnel endpoints are found at intersections of P- and Q-Spaces. In Figure 15-25, there is no common node that belongs to both reference areas and hence no viable repair tunnel endpoint is present.

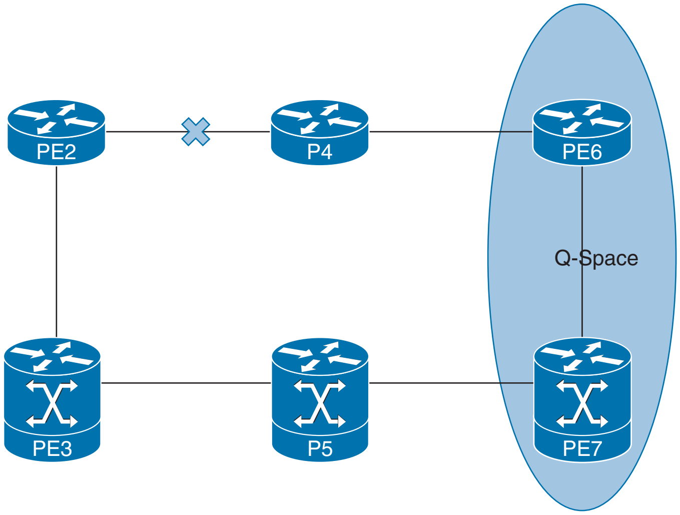


Figure 15-24 *Q-Space Reference Area*

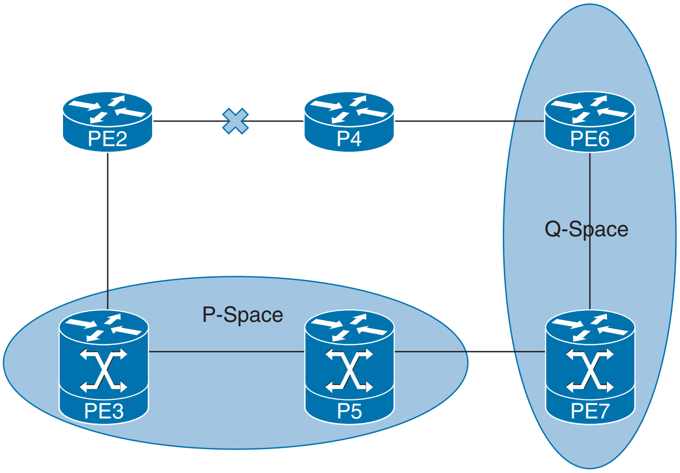


Figure 15-25 *P-Space and Q-Space Reference Area*

Extended P-Space

Because PE4 needs to repair the protected PE4–P2 link and reach any router in this ring topology without using the protected link, the concept of **Extended P-Space** was introduced. Extended P-Space is the union of each of PE4's neighbors. In this case, this is router PE3 in Figure 15-26, whose P-Space contains routers P5 and PE7. By combining P-Spaces of PE2 and PE3, we extend PE2's reach, and PE7 becomes a common point for P- and Q-Spaces. A **PQ node** of a node PE2, in relation to a protected link PE2–P4, is a node that belongs to both the P-space (or extended P-space) of PE2 for that link and the Q-space of P4 for the same link. PE7 is chosen as the repair tunnel endpoint. Why? Because repair tunnels are chosen from a set of PQ nodes.

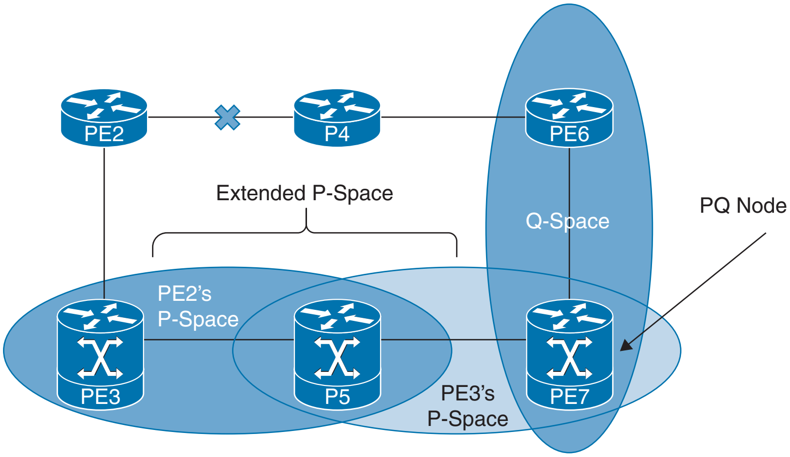


Figure 15-26 *PQ Node and Extended P-Space*

Classic LFA Limitations

Now, with the understanding of the reference points, Classic LFA's (loop-free alternate fast reroute, aka LFA-FRR) limitations become obvious. Note that I am not discussing LFA-FRR because it is not a part of the exam blueprint. I reference it here to highlight the advantages of TI-LFA. Figure 15-27 considers two such limitations.

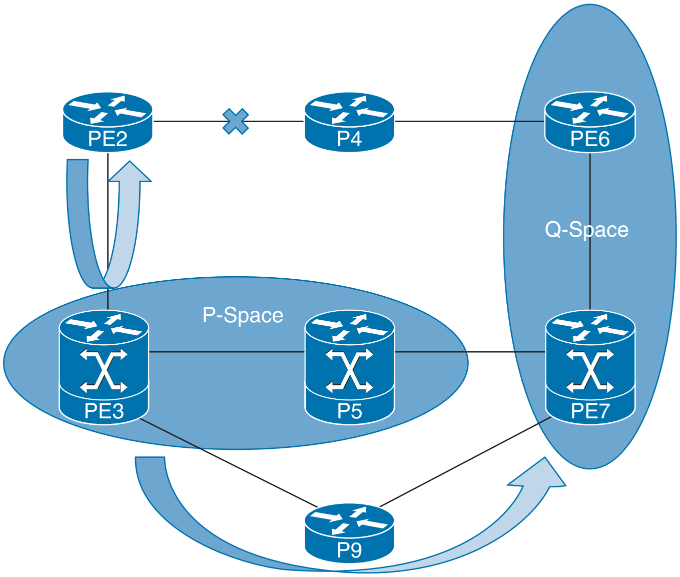


Figure 15-27 *Classic LFA Limitations Examples*

First, LFA-FRR suffers from incomplete coverage, which makes it topology dependent (as opposed to TI-LFA, which is topology independent). Recall our discussion about the PQ node. PE2 protects the PE2–P4 link and sends traffic to PE6. When the PE2–P4 link fails, PE2 will send traffic to PE3. Before the network converges via IGP, PE3 has a problem, since the shortest path to PE8 is still through the failed PE2–P4 link and PE3 will send the traffic back to PE2, looping the doomed packets. This is a real problem that, in the rLFA (Remote LFA) cases, can sometimes be solved by a Targeted LDP session, where PE2 would establish a remote LDP session with PE7, but this approach also has limitations that are outside of the scope of this exam. TI-LFA handles this topology though a “double-segment” coverage, where two labels are pushed (PE3, PE3-R5 ADJ-SID) to overcome this problem.

Second, notice the additional P9 router. Let’s suppose it is not a part of the network core or planned for capacity. Classic LFA will steer the traffic on this suboptimal backup path. Additional case-specific operator involvement would be necessary to avoid such undesired backup paths.

In contrast, a topology-independent loop free alternate (TI-LFA) provides 100 percent coverage and uses the post-convergence path as the fast reroute (FRR) backup path.

Key Topic

TI-LFA delivers significant improvements over the traditional loop-free alternate fast reroute (LFA-FRR) approach. TI-LFA uses a post-convergence path after a link failure occurs. This path is known before a failure occurs and is preprogrammed into the data plane. TI-LFA uses PQ nodes, or a combination of P and Q nodes located on the post-convergence path to compute backup paths. Traffic will be rerouted in sub-50ms on any topology.

While the blueprint does not focus on configuration of Segment Routing features, you need to know how to configure TI-LFA. So, here is your homework for this section: return to Example 15-3, which I took from a massive production lab we have within Cisco to show the latest technologies. Study it and locate the two highlighted commands that start with **fast-reroute**. I recommend you enable this on all provider facing links; they will provide “automagic” protection mechanisms for your entire network without having to build backup tunnels. Know that TI-LFA works seamlessly with Flex-Algo we discussed in the previous section.

NOTE One last thing that will not come up on the exam (as TI-LFA is positioned as a better alternative). Be aware that the TI-LFA concept is not new. RFC 4090 (*Fast Reroute Extensions to RSVP-TE for LSP Tunnels*) described this very technique in 2005, 10 years before Segment Routing hit the street) in Section 3.1 (One-to-one Backup method), albeit RSVP used signaling for tunnels and Segment Routing uses a label stack to guide the packet onto the new path. I’m here to assist you in preparing for your exam and provide valuable background information, without engaging in debates over differing viewpoints.

PCE-PCC Architecture

PCE-PCC architecture involves a Path Computation Element (PCE) that centrally computes optimal network paths and a Path Computation Client (PCC) that requests these paths, enabling efficient and scalable traffic engineering across the network. To take a step back, Segment Routing Traffic Engineering (SRTE) allows the network operator to force a packet anywhere on the network. The ingress router will contain the policy containing the operator’s intent. If the network is small like the basic topology we have been using for our examples, there are only a handful for routers that we have to individually configure with such policies. A great majority of service provider networks you are likely to encounter in your career will contain dozens, hundreds, or maybe thousands of nodes. The task of deploying a uniform policy on such a distributed network domain becomes laborious and operationally costly. How do you scale this type of rollout? There are many other limitations operators have encountered on distributed SR (or RSVP-TE for that matter) networks. Among the more notable ones is stale policies, in which operators define a set of policies and in six months traffic patterns change, which leads to continuous "rinse-and-repeat" of policy redeployment. Another one would be applications requesting the best available path in real time—not something that can by automatically done with the SRTE approach we have described thus far. What about being able to offer paths that meet certain SLAs? There are many other ones.

From the beginning of Traffic Engineering, the need for a centralized optimization element that can dynamically adjust policies based on current network conditions was apparent. Enter Path Computation Element Protocol (PCEP). It was initially specified to support the classic RSVP-TE protocol. With the introduction of Segment Routing, PCEP has been extended to support SRTE. RFC 4655 defines multiple terms that support PCEP-based architecture. Of immediate interest to us are the following terms:

- **Path Computation Element (PCE)**, which is "an entity that can compute a network path or route based on a network graph, and of applying computational constraints during the computation. The PCE entity is an application that can be located within a network node or component, on an out-of-network server, etc.…" (RFC 4655)
- **Path Computation Client (PCC)**, which is "a client application requesting a path computation to be performed by the Path Computation Element…" (RFC 4655)
- **Path Computation Element Protocol (PCEP)**, which is north-bound API capable, meaning that it can ingest information coming from the network routers (via BGP-LS updates, for example) and make real-time Traffic Engineering decisions based on current network conditions. This is extremely powerful and desired on modern networks.

Notice that you can run PCE on the router itself or can rely on another adjunct processor to perform this function. In the case of Cisco products, that would be the Crosswork Network Controller, which provides a wide assortment of functionalities that helps customers to simplify and automate intent-based network service provisioning, visualization, monitoring, and optimization in a multivendor network environment with a common GUI and API. In Cisco's documentation, you will often encounter references to SR-PCE. When you see these, it will either be a router running PCE or the Crosswork Network Controller. It can also think of PCE as a BGP Route-Reflector for Segment Routing and associated services. The following is a partial list of its capabilities (get the overall picture, do not memorize these for the exam):

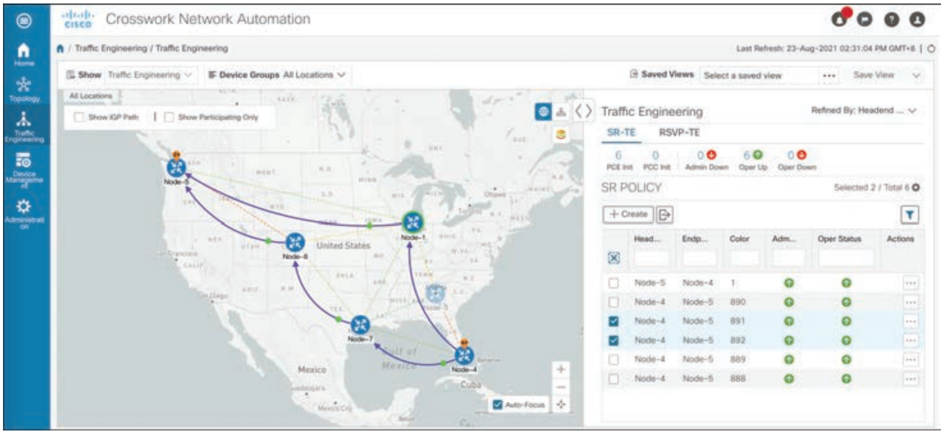
- Segment Routing (SR) policy provisioning with explicit intent (for example, bandwidth constraints, latency minimization, etc.).
- Services provisioning (for example, L2VPN, L3VPN services with associated segment routing policy).
- Collection of real-time performance information and network optimization to maintain the intent of the associated segment routing policy.
- Tactical optimization of the network during times of congestion.

- Assistance with migration to next-generation networks and technologies (for example, migration from RSVP-TE to SR-TE, implementing multicast with SR Tree-SID, embracing 5G network slicing, etc.).
- Monitoring and troubleshooting the health of L2VPN and L3VPN services through empirical data plane verification.
- Streamlining and automating network-focused Method of Procedure (MOP) for remediation and maintenance tasks.

Key Topic

What makes the SR-PCE controller so powerful is that it provided centralized SRTE visibility into multidomain topologies, something that SRTE routers are not able to deliver. Northbound APIs allow SR-PCE to compute paths in real time. Because of the above, the SR-PCE can construct SLA-aware path computations even across network domains while delivering end-to-end network topology awareness. Again, you should not view SR-PCE as a single all-overseeing device but rather think of a BGP Route-Reflector deployment model where intent is centrally disseminated.

Figure 15-28 shows a screenshot taken from the Crosswork Network Controller's GUI console.



☒	Head...	Endp...	Color	Adm...	Oper Status	Actions
□	Node-5	Node-4	1	●	●	···
□	Node-4	Node-5	890	●	●	···
☑	Node-4	Node-5	891	●	●	···
☑	Node-4	Node-5	892	●	●	···
□	Node-4	Node-5	889	●	●	···
□	Node-4	Node-5	888	●	●	···

Figure 15-28 *Crosswork Network Controller*

The Cisco Crosswork Optimization Engine stands as a key element within the Crosswork Automation Suite, offering real-time network optimization capabilities. Network operators can enhance network utility and accelerate service deployment through dynamic Traffic Engineering and proactive optimization. Working seamlessly with the Crosswork Optimization Engine, the WAN Automation Engine (WAE) caters to diverse aspects of capacity management. It spans from long-term network engineering to capacity planning and Traffic Engineering, ensuring optimal network operation under various conditions. Furthermore, the WAN Automation Engine serves a valuable role in simulation analysis, aiding in the identification of potential network hotspots during failure scenarios.

Exam Preparation Tasks

As mentioned in the section “How to Use This Book” in the Introduction, you have a few choices for exam preparation: the exercises here, Chapter 23, “Final Preparation,” and the exam simulation questions in the Pearson Test Prep Software Online.

Review All Key Topics

Review the most important topics in this chapter, noted with the Key Topic icon in the margin of the page. Table 15-5 lists a reference of these key topics and the page numbers on which each is found.

Key Topic

Table 15-5 Key Topics for Chapter 15

Key Topic Element	Description	Page Number
Table 15-2	LSD Label Ranges	729
Table 15-3	IS-IS TLVs	736
Example 15-3	Commands to Turn on Segment Routing in IS-IS	736
Table 15-4	OSPF Opaque LSAs	738
Figure 15-11	IPv6 Segment Routing Header Format	743
Figure 15-12	IPv6 Segment Identifier	743
Paragraph, Figure 15-17	Segment Routing Policy construction	749
Section	Binding-SID (BSID)	750
Paragraph	SRv6 encapsulation	753
Section	Terms from Remote LFA Technology	754
Paragraph	TI-LFA improvements	757
Paragraph	Crosswork considerations	759

Define Key Terms

Define the following key terms from this chapter and check the answers in the glossary:

BGP control plane, Binding Segment Identifier (BSID), Candidate Paths, Extended P-Space, Flex-Algo, Global Segments, IGP Adjacency Segment, IGP Prefix Segment, IS-IS Control Plane, Label Distribution Protocol (LDP), Label Switching Database (LSD), Local segment, OSPFv2 Control Plane, P-Space, Path Computation Client (PCC), Path Computation Element (PCE), Path Computation Element Protocol (PCEP), PCE-PCC architecture, PQ node, Q-space, Segment Routing, Segment Routing Control Plane, Segment Routing Global Block (SRGB), SR-MPLS (Segment Routing based on MPLS data plane), SRv6 (Segment Routing based on IPv6 data plane), SRv6 control plane, Topology-Independent Loop-free Alternate (TI-LFA)

Command Reference to Check Your Memory

This section includes the most important configuration and EXEC commands covered in this chapter. You might not need to memorize the complete syntax of every command, but you should be able to remember the basic keywords that are needed.

To test your memory of the commands, cover the right side of Table 15-6 with a piece of paper, read the description on the left side, and then see how much of the command you can remember.

The 350-501 exam focuses on practical, hands-on skills that are used by networking professionals. Therefore, you should be able to identify the commands needed to configure and test. Note that not all commands are fully covered in the chapter, but their presence in the table below should lead you to investigate them further to understand this technology.

Table 15-6 CLI Commands to Know

Task	Command Syntax
Define Segment Routing Global Block Range	RP/0/0/CPU0:P3(config)# **segment-routing global-block**
Configure IS-IS advertisements to BGP-LS	RP/0/0/CPU0:P3(config-isis)# **distribute link-state**
Configure IS-IS to generate and accept only new-style type-length-value (TLV) objects	RP/0/0/CPU0:P3(config-isis-af)# **metric-style wide** [**transition**] [**level** { **1** | **2** }]
Enable Segment Routing for IPv4 addresses with MPLS data plane	RP/0/0/CPU0:P3(config-isis-af)# **segment-routing mpls**
Enable topology-independent loop-free alternate (TI-LFA) path using the IP fast reroute (FRR) mechanism	RP/0/0/CPU0:P3(config-isis-if)# **fast-reroute per-prefix** RP/0/0/CPU0:P3 (config-isis-if)# **fast-reroute per-prefix ti-lfa**
Configure the Segment Routing Mapping Server (SRMS)	RP/0/0/CPU0:P3(config)# **segment-routing mapping-server prefix-sid-map address-family ipv4 10.1.100.4/32 17000 range 100**
Trace the routes to a destination in a Segment Routing network	RP/0/0/CPU0:P3# **traceroute sr-mpls 10.1.100.2/32**
Set the preference of Segment Routing (SR) labels over Label Distribution Protocol (LDP) labels	RP/0/0/CPU0:P3(config-isis-af)# **segment-routing mpls** [**sr-prefer**]
Specify or advertise the prefix (node) segment ID (SID) as an index value in IS-IS	RP/0/0/CPU0:P3(config)# **router isis 100** RP/0/0/CPU0:P3(config-isis)# **interface loopback0** RP/0/0/CPU0:P3(config-isis-if)# **address-family ipv4 unicast** RP/0/0/CPU0:P3(config-isis-if-af)# **prefix-sid index 3**
Specify or advertise the prefix (node) segment ID (SID) as an absolute value in OSPF	RP/0/0/CPU0:P3# **configure** RP/0/0/CPU0:P3(config)# **router ospf 1** RP/0/0/CPU0:P3(config-ospf)# **area 0** RP/0/0/CPU0:P3(config-ospf-ar)# **interface loopback0** RP/0/0/CPU0:P3(config-ospf-ar-if)# **prefix-sid absolute 16003**

Task	Command Syntax
Specify the Binding SID (BSID) allocation behavior	RP/0/0/CPU0:P3# **configure** RP/0/0/CPU0:P3(config)# **segment-routing** RP/0/0/CPU0:P3(config-sr)# **traffic-eng** RP/0/0/CPU0:P3(config-sr-te)# **binding-sid explicit fallback-dynamic** RP/0/0/CPU0:P3(config-sr-te)# **policy SAMPLE** RP/0/0/CPU0:P3(config-sr-te-policy)# **binding-sid mpls 1000**
Configure SRv6-TE locator and Binding SID (BSID) behavior	RP/0/0/CPU0:P3# **configure** RP/0/0/CPU0:P3(config)# **segment-routing traffic-eng** RP/0/0/CPU0:P3(config-sr-te)# **srv6 locator loc1 binding-sid dynamic behavior ub6-encaps-reduced**
Globally enable SRv6	RP/0/0/CPU0:P3(config)# **segment-routing srv6**
Configure the SRv6 Locator	RP/0/0/CPU0:P3(config-srv6)# **locators** RP/0/0/CPU0:P3(config-srv6-locators)# **locator myLoc1** RP/0/0/CPU0:P3(config-srv6-locator)# **micro-segment behavior unode psp-usd** RP/0/0/CPU0:P3(config-srv6-locator)# **prefix 2001:0:8::/48**

Review Questions

As a part of the review, we encourage you to provide *a single-sentence answer* (keep your answers as short as possible) to the following questions. If you struggle to complete this answer in a single sentence, this may indicate a lack of clarity or reveal gaps in your understanding. We have constructed these questions to help you consolidate this chapter's information and extract the essence of the covered content.

The answers to these questions appear in Appendix A. For more practice with exam format questions, use the Pearson Test Prep Software Online.

1. How does the implementation of Segment Routing enhance network scalability and simplify Traffic Engineering compared to traditional routing protocols?
2. In what ways can Segment Routing contribute to improved network resiliency and faster convergence times, especially in the face of dynamic changes or failures?
3. How can Segment Routing adapt to support emerging trends such as 5G networks, edge computing, and the increasing demand for network automation?
4. What specific scenarios or network topologies benefit the most from using the TI-LFA loop-free backup path mechanism?
5. Can you name one key benefit of integrating SRv6 to meet the evolving demands of modern applications, services, and emerging technologies?

Bibliography

S. Bryant, C. Filsfils, S. Previdi, M. Shand, and N. So. RFC 7490, *Remote Loop-Free Alternate (LFA) Fast Reroute (FRR)*, https://www.ietf.org/rfc/rfc7490.txt, IETF, April 2015.

P. Camarillo, Ed. RFC 8986, *Segment Routing over IPv6 (SRv6) Network Programming*, https://www.ietf.org/rfc/rfc8986.txt, IETF, February 2021.

D. Dukes, Ed. RFC 8754, *IPv6 Segment Routing Header (SRH)*, https://www.ietf.org/rfc/rfc8754.txt, IETF, March 2020.

A. Farrel, J.-P. Vasseur, and J. Ash. RFC 4655, *A Path Computation Element (PCE)-Based Architecture*, https://www.ietf.org/rfc/rfc4655.txt, IETF, August 2006.

C. Filsfils. *Segment Routing, Part II: Traffic Engineering*, Self-published, 2019 (ISBN: 978-1095963135).

C. Filsfils, K. Talaulikar, Ed., D. Voyer, A. Bogdanov, and P. Mattes. RFC 9256, *Segment Routing Policy Architecture*, https://www.ietf.org/rfc/rfc9256.txt IETF, July 2022.

L. Ginsberg, Ed. RFC 8667, *IS-IS Extensions for Segment Routing*, https://www.ietf.org/rfc/rfc8667.txt, IETF, December 2019.

LabN. RFC 5250, *The OSPF Opaque LSA Option*, https://www.ietf.org/rfc/rfc5250.txt, IETF, July 2008.

J. Liste. *A Guide to a Successful Segment Routing Deployment*, Cisco Live Presentation, 2023.

S. Litkowski, A. Bashandy, C. Filsfils, P. Francois, and B. Decraene. Internet Draft, *Topology Independent Fast Reroute Using Segment Routing*, https://www.ietf.org/archive/id/draft-ietf-rtgwg-segment-routing-ti-lfa-11.html, IETF, June 2023.

P. Pan, G. Swallow, and A. Atlast, Eds. RFC 4090, *Fast Reroute Extensions to RSVP-TE for LSP Tunnels*, https://www.ietf.org/rfc/rfc4090.txt, IETF, May 2005.

S. Previdi, Ed. RFC 8665, *OSPF Extensions for Segment Routing*, https://www.ietf.org/rfc/rfc8665.txt, IETF, December 2019.

S. Previdi. RFC 8670, *BGP Prefix Segment in Large-Scale Data Centers*, https://www.ietf.org/rfc/rfc8670.txt, IETF, December 2019.

S. Previdi and L. Ginsberg, Eds. RFC 8402, *Segment Routing Architecture*, https://www.ietf.org/rfc/rfc8402.txt, IETF, July 2018.

Redback Networks, Inc. RFC 5305, *IS-IS Extensions for Traffic Engineering*, https://www.ietf.org/rfc/rfc5305.txt, IETF, October 2008.

E. Rosen. RFC 8277, *Using BGP to Mind MPLS Labels to Address Prefixes*, https://www.ietf.org/rfc/rfc8277.txt, IETF, October 2017.

A. Roy. Internet Draft, *OSPFv3 LSA Extendibility*, https://datatracker.ietf.org/doc/html/draft-ietf-ospf-ospfv3-lsa-extend-23, IETF, January 2018.

Segment Routing website: https://www.segment-routing.net

This page intentionally left blank

Index