



## **Segment Routing Configuration Guide, Cisco IOS XE 17 (Cisco NCS 4200 Series)**

**First Published:** 2019-12-23

**Last Modified:** 2021-03-31

### **Americas Headquarters**

Cisco Systems, Inc.  
170 West Tasman Drive  
San Jose, CA 95134-1706  
USA  
<http://www.cisco.com>  
Tel: 408 526-4000  
800 553-NETS (6387)  
Fax: 408 527-0883

THE SPECIFICATIONS AND INFORMATION REGARDING THE PRODUCTS IN THIS MANUAL ARE SUBJECT TO CHANGE WITHOUT NOTICE. ALL STATEMENTS, INFORMATION, AND RECOMMENDATIONS IN THIS MANUAL ARE BELIEVED TO BE ACCURATE BUT ARE PRESENTED WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED. USERS MUST TAKE FULL RESPONSIBILITY FOR THEIR APPLICATION OF ANY PRODUCTS.

THE SOFTWARE LICENSE AND LIMITED WARRANTY FOR THE ACCOMPANYING PRODUCT ARE SET FORTH IN THE INFORMATION PACKET THAT SHIPPED WITH THE PRODUCT AND ARE INCORPORATED HEREIN BY THIS REFERENCE. IF YOU ARE UNABLE TO LOCATE THE SOFTWARE LICENSE OR LIMITED WARRANTY, CONTACT YOUR CISCO REPRESENTATIVE FOR A COPY.

The Cisco implementation of TCP header compression is an adaptation of a program developed by the University of California, Berkeley (UCB) as part of UCB's public domain version of the UNIX operating system. All rights reserved. Copyright © 1981, Regents of the University of California.

NOTWITHSTANDING ANY OTHER WARRANTY HEREIN, ALL DOCUMENT FILES AND SOFTWARE OF THESE SUPPLIERS ARE PROVIDED "AS IS" WITH ALL FAULTS. CISCO AND THE ABOVE-NAMED SUPPLIERS DISCLAIM ALL WARRANTIES, EXPRESSED OR IMPLIED, INCLUDING, WITHOUT LIMITATION, THOSE OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT OR ARISING FROM A COURSE OF DEALING, USAGE, OR TRADE PRACTICE.

IN NO EVENT SHALL CISCO OR ITS SUPPLIERS BE LIABLE FOR ANY INDIRECT, SPECIAL, CONSEQUENTIAL, OR INCIDENTAL DAMAGES, INCLUDING, WITHOUT LIMITATION, LOST PROFITS OR LOSS OR DAMAGE TO DATA ARISING OUT OF THE USE OR INABILITY TO USE THIS MANUAL, EVEN IF CISCO OR ITS SUPPLIERS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Any Internet Protocol (IP) addresses and phone numbers used in this document are not intended to be actual addresses and phone numbers. Any examples, command display output, network topology diagrams, and other figures included in the document are shown for illustrative purposes only. Any use of actual IP addresses or phone numbers in illustrative content is unintentional and coincidental.

All printed copies and duplicate soft copies of this document are considered uncontrolled. See the current online version for the latest version.

Cisco has more than 200 offices worldwide. Addresses and phone numbers are listed on the Cisco website at [www.cisco.com/go/offices](http://www.cisco.com/go/offices).

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. To view a list of Cisco trademarks, go to this URL: <https://www.cisco.com/c/en/us/about/legal/trademarks.html>. Third-party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1721R)

© 2019–2021 Cisco Systems, Inc. All rights reserved.



# CONTENTS

---

## CHAPTER 1

---

### Feature History 1

---

## CHAPTER 2

### Segment Routing Overview 9

How Does Segment Routing Work? 9

Examples for Segment Routing 10

Benefits of Segment Routing 12

Segment Routing Limitations 15

General Limitations of Segment Routing 16

Configuring Segment Routing 16

Configuring Segment Routing on an IGP Instance 17

Enabling Advertisement of Mapping Server Prefix Ranges 17

Disabling the Mapping Server 17

---

## CHAPTER 3

### Segment Routing Global Block (SRGB) 19

SRGB Limitations 19

Configuring the SRGB 19

Reserving a New SRGB 20

Modifying an Existing SRGB 20

Adjacency Segment Identifiers 20

Prefix Segment Identifiers 20

Types of Prefix SID Mapping 21

Export Prefix SID Mapping 21

Connected Prefix SID Mappings 21

Configuring Export Prefix SID Mapping 21

Configuring Connected Prefix SID Mapping 21

Configuring the Preferred Label 22

Configuring the Segment Routing Explicit Null	22
---	----

---

**CHAPTER 4**
**Using Segment Routing with IS-IS 23**

Restrictions for Using Segment Routing with IS-IS	23
Enabling Segment Routing	23
Enabling Segment Routing for IGPs	24
Configuring Segment Routing on IS-IS	24
Prefix-SID Received in LSPs from Remote routers	25
Limitations	25
Segment Routing Adjacency SID Advertisement	25
Multiple Adjacency-SIDs	26
Segment Routing Mapping Server (SRMS)	26
Connected Prefix SIDs	26
Configuring IS-IS SRMS	26
Configuring IS-IS SRMS Client	26
Configuring ISIS SID Binding TLV Domain Flooding	27
SRGB Range Changes	27
SRGB Deletion	27
MPLS Forwarding on an Interface	28
Segment Routing and LDP Preference	28
Segment Routing-TE	28
Enabling and Disabling SR-TE Announcements	28
RLFA LDP and SR	28
Topology-Independent LFA	29
Restrictions for the TI-LFA	30
Tie-breaker	30
Interface FRR Tiebreakers	31
Limitations on Tie-Beakers	32
Configuring T1 LFA	32
Configuration Example	33
Verifying the Tie-breaker	34
Verifying the Primary and Repair Paths	35
Verifying the IS-IS Segment Routing Configuration	36
Verifying the IS-IS TI-LFA Tunnels	37

**CHAPTER 5****Using Segment Routing with OSPF 39**

- Restrictions for Using Segment Routing with OSPF 39
- Enabling Segment Routing for OSPF 40
  - Configuring SR on an OSPF Instance 40
  - Enabling Advertisement of Mapping Server Prefix Ranges 40
  - Disabling Mapping Server Functionality 40
  - Interface Mode Prefix Attributes Command 41
- Using OSPF Fast Reroute with Segment Routing 41
  - Per Instance Enablement 42
  - Limitations of TI-LFA on OSPF 42
  - Enabling and Disabling TI-FRR 43
  - Verifying TI-LFA on OSPF 43
- Verifying the Configuration of SR on OSPF 44
- Troubleshooting Segment Routing on OSPF 47
- SR with OSPF Point-to-Multipoint Non Broadcast Mode 47
  - License Requirement 48
  - Limitation 48
  - Configure SR with OSPF P2MP NB Network 48
  - Verification of SR with OSPF P2MP NB Network Configuration 49
- Additional References 49

**CHAPTER 6****Segment Routing Traffic Engineering 51**

- Restrictions for SR-TE 52
- Segment Routing Recirculation for TI-LFA 53
  - Restrictions 53
  - Support for Seven SR-TE Lables 54
- Statistics Counters on Cisco RSP3 55
- Configuring SR-TE Tunnels 55
- Verifying SR-TE Tunnels 56
- Verifying Statistics Counters 56
- Configuring a Path Option for a TE Tunnel 57
- Configuring SR Explicit Path Hops 57
- Use Case: Segment Routing Traffic Engineering Basic Configuration 58

Explicit Path SR-TE Tunnel 1	59
Explicit Path SR-TE Tunnel 2	59
Explicit Path SR-TE Tunnel 3	60
Dynamic Path SR-TE Tunnel 4	60
Dynamic Path SR-TE Tunnel 5	60
Verifying Configuration of the SR-TE Tunnels	60
Verifying Tunnel 1	61
Verifying Tunnel 2	61
Verifying Tunnel 3	62
Verifying Tunnel 4	62
Verifying Tunnel 5	63
SR-TE LSP Instantiation	63
SR-TE LSP Explicit Null	64
SR-TE LSP Path Verification	64
Topology Path Validation	64
SR SID Validation	65
LSP Egress Interface	65
IP Reachability Validation	65
Tunnel Path Affinity Validation	65
Tunnel Path Resource Avoidance Validation	67
Tunnel Path Loop Validation	67
SR-TE Traffic Load Balancing	68
Load Balancing on Port Channel TE Links	68
Load Balancing on ECMPs	68
Load Balancing on Multiple Tunnels	68
SR-TE Tunnel Re-optimization	68
SR-TE With lockdown Option	69
SR-TE Tunnel Protection	70
IP-FRR Local Repair Protection	70
Tunnel Path Protection	71
SR-TE and TI-LFA	71
Restrictions for Using SR-TE and TI-LFA	71
Verifying the SR-TE With TI_LFA Configuration	73
Configuring TI-LFA With Mapping Server	75

<b>CHAPTER 7</b>	<b>SR-TE Per-Flow (Class) ODN and Automated Steering (PCE Delegated)</b>	<b>79</b>
	Restrictions for SR-TE Per-Flow (Class) ODN and Automated Steering (PCE Delegated)	79
	Configuring SR ODN Per-Flow Policy (PFP) AUTO STEERING (PCE DELEGATED)	81
<b>CHAPTER 8</b>	<b>SR-TE On Demand LSP</b>	<b>83</b>
	Restrictions for SR-TE On Demand LSP	83
	Information About SR-TE On Demand LSP	83
	SR-TE: Setup LSP as Static Route	83
	Static SRTE over Unnumbered Interfaces	84
	How to Configure SR-TE On Demand LSP	84
	Configuring LSP as Static Route	84
	Enabling Segment Routing Auto Tunnel Static Route	85
	Verifying Segment Routing Auto-Tunnel Static Route	85
<b>CHAPTER 9</b>	<b>SR-TE On-Demand Next Hop</b>	<b>89</b>
	Restrictions for SR-TE On-Demand Next Hop	89
	Information About SR-TE On-Demand Next Hop	89
	Fast Convergence Default Optimize	90
	Default Optimize Values for IS-IS	91
	Default Optimize Values for OSPF	91
	SR-TE Policy, Color Extended Community, Affinity Constraint, and Disjointness Constraint	93
	SR-TE Policy Command	94
	Color Extended Community	94
	Affinity Constraint	94
	Disjointness Constraint	94
	Limitations	95
	How to Configure Segment Routing On-Demand Next Hops—Layer 3 and Layer 3 VPN	95
	Configuring Segment Routing On-Demand Next Hops—Layer 3 and Layer 3 VPN	95
	Verifying Segment Routing On-Demand Next Hop—Layer 3 and Layer 3 VPN	99
	Configuring Color Extended Community, Affinity Constraint, and Disjointness Constraint	103
	Configuring Color Extended Community	104
	Configuring Affinity Constraint	105
	Configuring Disjointness Constraint	105

Verifying SR-TE ODN Color Extended Community, Affinity Constraint, and Disjointness Constraint	106
Troubleshooting the SR-TE ODN Color Extended Community, Affinity Constraint, and Disjointness Constraint	110

---

**CHAPTER 10**
**Segment Routing OAM Support 115**

Restrictions for Segment Routing OAM Support	115
Information About Segment Routing-OAM Support	115
Segment Routing-OAM Support	115
LSP Ping Operation for Nil FEC target	116
How to Diagnose Segment Routing with LSP Ping and Trace Route Nil FEC Target	116
Use LSP Ping for Nil FEC Target	116
Using LSP Traceroute for Nil FEC Target	117
Example: LSP Ping Nil_FEC Target Support	117

---

**CHAPTER 11**
**Dynamic Path Computation Client 121**

Information About Dynamic PCC	121
Configuring Dynamic PCC Globally	122
Configuring Dynamic PCC on an Interface	122
Verifying Dynamic PCC	122

---

**CHAPTER 12**
**Segment Routing uLoop Avoidance 127**

Prerequisites for Segment Routing uLoop Avoidance	127
Restrictions for Segment Routing uLoop Avoidance	127
Information About Segment Routing uLoop Avoidance	127
Microloops	127
How Segment Routing Prevents Microloops?	130
How to Enable Segment Routing uLoop Avoidance	131
Enabling ISIS - Microloop Avoidance	131
Verifying ISIS - Microloop Avoidance	131
Enabling OSPF - Microloop Avoidance	132
Verifying OSPF - Microloop Avoidance	133

---

**CHAPTER 13**
**EVPN-VPWS Single Homing on Segment Routing 135**

Information About EVPN-VPWS	135
-----------------------------	-----



Benefits of EVPN-VPWS Single Homed	136
Prerequisites for EVPN-VPWS	136
Restrictions for EVPN-VPWS	136
How to Configure EPVN-VPWS	137
Configuring BGP for EVPN-VPWS	137
Configuring EVPN-VPWS Instance	137
Rewrite for EVI Service Instance	137
Configuring EVPN-VPWS for Logging	138
Verifying EVPN-VPWS Instance	138
Verifying EVPN-VPWS Configuration	138
Verifying EVPN-VPWS Configuration for Logging	140
Troubleshooting	141
Virtual Circuit (VC) is in Down state	141
VC FSM History	142
Remote-Wait State	143
Configuration Examples for EVPN-VPWS over SR	143
Additional References for EVPN-VPWS	147

---

**CHAPTER 14**
**Configure Performance Measurement 149**

Link Delay Measurement	150
Restrictions and Usage Guidelines for PM for Link Delay	151
PM Link Delay: Default Values for Different Parameters	151
Configuration Example: PM for Link Delay	152
Verification: PM Link Delay Configuration	152
End-to-End Delay Measurement	154
Configuration Example: PM for End-to-End Delay Management	155
Verification: PM End-to-End Delay Management Configuration	156
SR-PCE: Enabling SR PM Delay or Liveness for PCE-Initiated Policies	158
COE-PCE Initiated SR Policy	159
Configure SR-PCE: Enabling SR-PM Delay or Liveness for PCE-Initiated Policies	159
Verification of SR-PCE: Enabling SR-PM Delay or Liveness for PCE-Initiated Policies	161
Telemetry (Model-Based Telemetry and Event-Based Telemetry) Support for Performance Measurement	162
Configuration Example: Telemetry for Performance Measurement	164

Verification of MDT and EDT Support for Performance Measurement	165
Configuring UDP Destination Port	168

---

<b>CHAPTER 15</b>	<b>Static Route Traffic Steering Using SR-TE Policy</b>	<b>171</b>
	Restrictions	172
	Configuration Example: Static Route Traffic Steering Using SR-TE Policy	172
	Verification of Static Route Traffic Steering Using SR-TE Configuration	173

---

<b>CHAPTER 16</b>	<b>EVPN Single-Homing Over Segment Routing</b>	<b>175</b>
	Information about EVPN Single-Homing	176
	Ethernet Multipoint Connectivity	176
	EVPN Multipoint Solution	176
	EVPN Building Blocks	176
	Service Interfaces	176
	Route Types	178
	Prerequisites for EVPN Single-Homing	180
	Restrictions for EVPN Single-Homing	180
	How to Configure EVPN Single Homing	181
	Configuring EVPN	181
	Configuring EVPN Single-Homing	183
	Verification Examples for EVPN Single-Homing	185
	Additional References for EVPN Single-Homing	191

---

<b>CHAPTER 17</b>	<b>Enabling Segment Routing Flexible Algorithm</b>	<b>193</b>
	Prerequisites for Flexible Algorithm	195
	Restrictions for Flexible Algorithm	195
	Building Blocks of Segment Routing Flexible Algorithm	195
	Flexible Algorithm Definition	195
	Flexible Algorithm Support Advertisement	195
	Flexible Algorithm Definition Advertisement	195
	Flexible Algorithm Prefix-SID Advertisement	196
	Calculation of Flexible Algorithm Path	196
	OSPF Flexible Algorithm TI-LFA	197
	OSPF Flexible Algorithm TI-LFA Computation	197

Configuring OSPF Flexible Algorithm TI-LFA	198
Verifying the OSPF Flexible Algorithm TI-LFA	199
Installation of Forwarding Entries for Flexible Algorithm Paths	201
Flexible Algorithm Configurations	201
Configuring IS-IS Flexible Algorithm	203
Configuring OSPF Flexible Algorithm	204
Configuring Prefix-SID	205
Configuring SRTE-ODN Assosiation	205
Configuring the Interface for Flexible Algorithm	205
Configuring BGP	205
Configuring Selective Path Filtering	206
Verifying the Flexible Algorithm Configuration	207

---

## CHAPTER 18

### COE-PCE Initiated SR Policy with OSPF and IS-IS SR-TE Autoroute Announce 219

COE-PCE Initiated SR Policy with OSPF Autoroute Announce	219
COE-PCE Initiated SR Policy	220
Restrictions for PCE Initiated SR Policy	220
ECMP Over SR-TE	221
Restrictions for ECMP over SR-TE Policies	221
Local Congestion Mitigation	222
Load Balancing	223
Configuring ECMP over SR-TE Policy with OSPF Autoroute Announce	224
SR-PCE: Support for PCE-Initiated SR Policy and ISIS Autoroute	226
Configure PCE-Initiated SR Policy and ISIS Autoroute	226
Verification of IS-IS Autoroute Configuration	227

---

## CHAPTER 19

### L2VPN over SR-TE Preferred Path 229

Restrictions	230
Configuring L2VPN over SR-TE Preferred Path	230
Configuration Example: L2VPN over SR-TE Preferred Path	231
Verification of L2VPN over SR-TE Preferred Path Configuration	232

---

## CHAPTER 20

### SR-TE PM: Liveness of SR Policy Endpoint 235

Benefits	236
----------	-----

Restrictions	236
Probe Query Message	237
SR-TE Re-optimization (Make-Before-Break)	237
Performance Measurement Named Profile	237
Configure SR-TE PM: Liveness of SR Policy Endpoint	237
Configuration Example: Default Delay Profile for Liveness	238
Configuration Example: PM Named Profiles	238
Configuration Example: SR Policy for Liveness Detection	239
Verification of SR-TE PM: Liveness of SR Policy Endpoint Configuration	240
SR-PM Delay Deduction (Loopback Mode)	245
Configuration Example: SR-PM Delay Deduction (Loopback Mode)	246
Verify SR-PM Delay Deduction (Loopback Mode) Configuration	246

---

## CHAPTER 21

### **EVPN Integrated Routing and Bridging (L2 and L3 Anycast Gateway) and Data Center Interconnect or Border Leaf (Single Homing) 251**

Distributed Anycast Gateway (DAG) with Bridge Domains	252
Symmetric IRB with MPLS on Distributed Gateways	253
Host MAC-IP Binding on a Single-Homed DAG	254
Host MAC-IP Mobility	254
ARP and ND Flooding Suppression	255
Prerequisites	255
Restrictions	256
Configuring EVPN Integrated Routing and Bridging (L2 and L3 Anycast Gateway) and Data Center Interconnect or Border Leaf (Single Homing)	256
Configure IP VRF	256
Configure Layer 2 Virtual Private Network EVPN	257
Configure Bridge Domain	257
Configure Bridge Domain IRB Interface	257
Configure BGP IRB	257
Verification of EVPN Integrated Routing and Bridging (L2 and L3 Anycast Gateway) and Data Center Interconnect or Border Leaf (Single Homing) Configuration	258
Verify Device Tracking Database	258
Verify L2VPN EVPN Summary	259
Verify L2VPN EVPN EVI	260
Verify Platform Software Infrastructure Punt Statistics	260

Verify Platform Software Infrastructure Inject	261
Verify BGP L2VPN EVPN Detail	262
EVPN-IRB DHCP v4 and v6 Relay over Segment Routing	264
How DHCP Relay Agent Works	265
Enhancement of Flooding Suppression Feature	266
Restrictions	267
Configure DHCP Relay Flooding Suppression	267
Configure DHCP Relay on DAGs	267
Verification of DHCPv4 and DHCPv6 Relay Configuration	268
Stitching of Subnet Route from EVPN to L3VPN	271
Restrictions	273
Scale	273
Configure BGP L3VPN Subnet Advertisement (Stitching) for Collapsed Spine or Leaf or RR on a Single Node	274
Verification of BGP L3VPN Subnet Advertisement (Stitching) for Collapsed Spine or Leaf or RR on a Single Node Configuration	275





## CHAPTER 1

# Feature History

The following table lists the new and modified features supported in the Segment Routing Configuration Guide in Cisco IOS XE 17 releases, on Cisco NCS 4201 and Cisco NCS 4202 routers.

Feature	Description
<b>Cisco IOS XE Cupertino 17.7.1</b>	
<a href="#">SR with OSPF Point-to-Multipoint Non Broadcast Mode</a>	This feature enables SR on Point-to-Multipoint (P2MP) Non Broadcast (NB) mode, where you can manually configure neighbor IP address and OSPF sends hello packet using the neighbor unicast address.
<a href="#">SR-PCE: Support for PCE-Initiated SR Policy and ISIS Autoroute</a>	This feature enables System-to-Intermediate System (IS-IS) interaction with traffic engineering to receive the SR-TE policies via autoroute announcement notification. These notifications are used as IGP shortcuts during SPT computation and route calculation and are installed as nexthops for applicable routes in Routing Information Base (RIB) or MPLS Forwarding Infrastructure (MFI).
<b>Cisco IOS XE Bengaluru 17.6.1</b>	
<a href="#">EVPN-IRB DHCP v4 and v6 Relay over Segment Routing</a>	<p>This feature introduces a specialised implementation of DHCP packets to support DHCPv4 and DHCPv6 in an EVPN Fabric with Distributed Anycast Gateways (DAGs) on the same Virtual Routing and Forwarding (VRF). It also avoids DHCP discovery packet floods across the fabric.</p> <p>The flooding suppression feature is also enhanced to intercept multicast or broadcast DHCP packets when DHCP relay is configured on the DAG to perform the required action and localize the scope of the service.</p>
<a href="#">IS-IS Flexible Algorithm Include Affinity Support</a>	This feature supports "include-any" and "include-all" affinities in IS-IS. Prior to Cisco IOS XE Bengaluru 17.6.1 release, only Flexible Algorithm affinity "exclude-any" was supported.
<a href="#">OSPF Flexible Algorithm (Ph2): Topology-Independent Loop-Free Alternate (TI-LFA) path</a>	This feature allows you to configure the Loop-Free Alternate (LFA) and TI-LFA backup or repair paths for a Flexible Algorithm. The backup path is computed based on the constraints and metrics of the primary path. Prior to Cisco IOS XE Bengaluru 17.6.1, OSPF Flexible Algorithm supported only the primary path.

Feature	Description
<a href="#">SR-PCE: Enabling SR PM Delay or Liveness for PCE-Initiated Policies</a>	This feature enables the Path Computation Element (PCE) that can provision a Segment Routing Traffic Engineering (SR-TE) policy to mitigate link congestion. Prior to this release, you could only enable PM link and delay measurement using CLI-based policies. Starting with this release, you can also use PCE to enable PM link and delay measurement.
<a href="#">Stitching of Subnet Route from EVPN to L3VPN</a>	This feature introduces the collapsed spine and border leaf node in the network topology of single homing DAGs with symmetric IRB, inter-subnet layer 3 traffic within fabric and inter-subnet layer 3 stitching through layer 3 border gateway. The hosts participating in fabric IRB are directly attached with the collapsed spine and border leaf node.
<b>Cisco IOS XE Bengaluru 17.5.1</b>	
<a href="#">SR-PM Delay Deduction (Loopback Mode)</a>	This feature improves the SR-PM detection time as the PM probes are not punted on the remote nodes. Also, it does not require a third-party support for interoperability.
<a href="#">SR-TE PM: Liveness of SR Policy Endpoint</a>	This feature enables Performance Measurement (PM) liveness detection and delay measurement for an SR policy on all the segment lists of every candidate path that are present in the forwarding table using PM probes. Thus, you can easily monitor the traffic path and efficiently detect any drop of traffic due to cable or hardware or configuration failures. This feature provides the following benefits: <ul style="list-style-type: none"> <li>• End-to-end liveness is verified before activating the candidate path in the forwarding table.</li> <li>• End-to-end liveness failure can trigger re-optimization to another path by deactivating the current path.</li> </ul>
<a href="#">Segment Routing Flexible Algorithm with OSPF</a>	This feature allows you to configure Segment Routing Flexible Algorithm with OSPF. Flexible Algorithm with OSPF supports metric minimization and avoidance, multi-plane, delay metric with rounding, and ODN with auto-steering.
<a href="#">Segment Routing Policy Counters</a>	This feature enables statistic counters to be displayed when traffic passes over the SR-TE tunnel.  You can use the command <b>show segment-routing traffic-eng policy name</b> <i>policy name</i> to view the counters.
<b>Cisco IOS XE Bengaluru 17.4.1</b>	
<a href="#">L2VPN over SR-TE Preferred Path</a>	This feature allows you to configure an SR policy as the preferred path for a VPWS or VPLS pseudowire. VPWS or VPLS pseudowires between same PEs can be routed over different SR policies based on the requirements. Prior to this release, you could only steer the traffic using the SR policy for routing IPv4 traffic to a destination pseudowire (over IGP or BGP-LU).
<a href="#">PCE Initiated SR Policy with OSPF Autoroute Announce</a>	This feature enables a steering mechanism in which IGPs automatically use the policy for destination's downstream of the policy end point.



Feature	Description
<a href="#">Segment Routing Flexible Algorithm support for TI-LFA uLoop Avoidance, SID Leaking, and ODN with Auto-Steering</a>	<p>This feature allows you to compute Loop Free Alternate (LFA) paths, TI-LFA backup paths, and Microloop Avoidance paths for a particular Flexible Algorithm using the same constraints as the calculation of the primary paths for such Flexible Algorithms, for IS-IS. See <a href="#">Calculation of Flexible Algorithm Path</a>.</p> <p>Inter-area leaking of Flexible Algorithm SIDs and prefixes and selectively filtering the paths that are installed to the MFI are also supported. See <a href="#">Flexible Algorithm Prefix-SID Advertisement</a> and <a href="#">Installation of Forwarding Entries for Flexible Algorithm Paths</a>.</p>
<a href="#">Telemetry (Model-Based Telemetry and Event-Based Telemetry) Support for Performance Measurement</a>	This feature enables Model-Based Telemetry (MDT) and Event-Based Telemetry (EDT) that allow the data to be directed to a configured receiver. This data can be used for analysis and troubleshooting purposes to maintain the health of the network. The <code>sr_5_label_push_enable</code> SDM template is mandatory for this feature to function.
<b>Cisco IOS XE Amsterdam 17.3.1</b>	
<a href="#">EVPN Single-Homing Over Segment Routing</a>	The EVPN Single-Homing feature utilizes the BGP MPLS-based Ethernet VPN functionality as defined in RFC 7432. For EVPN Single-Homing feature, a CE device is attached to a single PE device and has an Ethernet Segment.
<a href="#">SR-TE Per-Flow (Class) ODN and Automated Steering (PCE Delegated)</a>	This feature lets you steer traffic with SR-TE PFP based on the QoS markings on the packets. The traffic is then switched onto the appropriate path based on the forward classes of the packet. This feature is supported on the Cisco RSP2 and RSP3 modules.
<a href="#">Segment Routing Performance Measurement Delay Measurement Using RFC 5357 (TWAMP Light)</a>	This feature enables hardware timestamping. The Performance Measurement (PM) for link delay uses the light version of Two-Way Active Measurement Protocol (TWAMP) over IP and UDP defined in Appendix I of RFC 5357. TWAMP provides an alternative for interoperability when RFC 6374 is not used.
<a href="#">Segment Routing Flexible Algorithm</a>	This feature allows you to customize IGP shortest path computation according to your needs. You can assign custom SR prefix-SIDs to forward the packets beyond link-cost-based SPF. As a result, a traffic engineered path is automatically computed by the IGP to any destination reachable by the IGP.
<a href="#">Segment Routing Performance Measurement End-to-End Delay Measurement</a>	This feature allows to monitor the end-to-end delay experienced by the traffic sent over a Segment Routing policy. This feature ensures the delay does not exceed the specified threshold value and violate the SLAs. Use this feature to apply extended TE link delay metric (minimum delay value) to compute paths for Segment Routing policies as an optimization metric or as an accumulated delay bound.

Feature	Description
<a href="#">Static Route Traffic Steering Using SR-TE Policy</a>	This feature allows the non colored (BGP Extended Community) prefix to steer traffic over static policy. Prior to this release, only colored (BGP Extended Community) prefix could automatically steer traffic based on the defined policy using a tunnel interface. Unlike non colored prefix, this was possible only for the colored prefix as it could match the SR policy. IPv4 static routes are now enhanced to leverage the SR policies to aid Segment Routing Traffic Engineering (SR-TE). This facilitates traffic steering for non colored prefix as you can now configure IP Static Route with SR static policy. The following new keyword for the <b>ip route</b> command is introduced: <b>segment-routing policy</b> [ <i>policy name</i> ].
<b>Cisco IOS XE Amsterdam 17.1.1</b>	
<a href="#">EVPN-VPWS Single Homing over Segment Routing</a>	EVPN-VPWS single homing is a BGP control plane solution for point-to-point services. It has the ability to forward traffic from one network to another using Ethernet Segment without MAC lookup. EVPN-VPWS single homing works on both IP and SR core. IP core is used to support BGP while the SR core is used to switch packets between the endpoints.
<a href="#">Segment Routing Low Latency Network Slice</a>	This feature allows the advertisement and reception of the extended TE link delay metrics without any additional configuration required in IS-IS, OSPF or BGP-IS. When the link delay values are configured, they are flooded in the PCE topology and when the path computation is requested, these values are used for path calculation.
<a href="#">Segment Routing Performance Measurement Link Delay Metrics</a>	Network performance metrics such as packet loss, delay, delay variation, and bandwidth utilization is a critical measure for traffic engineering (TE) in service provider networks. These metrics provide network operators with information about characteristics of their networks for performance evaluation and helps to ensure compliance with service level agreements. The service-level agreements (SLAs) of service providers depend on the ability to measure and monitor these network performance metrics.
<a href="#">SR-TE Policy for MPLS TE</a>	The Cisco ASR 900 Series routers with Cisco RSP2 module support the newer segment routed Traffic Engineering (SR-TE) policy and you can enable the policy using the segment-routing traffic-eng command.
<a href="#">SR-TE ODN Color Extended Community for Layer 3 VPN</a>	The Cisco ASR 900 Series routers with Cisco RSP2 module support the color extended community with the following feature support: <ul style="list-style-type: none"> <li>• An egress router adds the color extended community to the BGP updates that require a Traffic-Engineered path.</li> <li>• An SR-TE policy is created on the ingress router for the color-endpoint pair.</li> </ul>

The following table lists the new and modified features supported in the Segment Routing Configuration Guide in Cisco IOS XE 17 releases, on Cisco NCS 4206 and Cisco NCS 4216 routers.

Feature	Description
<b>Cisco IOS XE Cupertino 17.7.1</b>	
<a href="#">SR with OSPF Point-to-Multipoint Non Broadcast Mode</a>	This feature enables SR on Point-to-Multipoint (P2MP) Non Broadcast (NB) mode, where you can manually configure neighbor IP address and OSPF sends hello packet using the neighbor unicast address.

Feature	Description
<a href="#">SR-PCE: Support for PCE-Initiated SR Policy and ISIS Autoroute</a>	This feature enables System-to-Intermediate System (IS-IS) interaction with traffic engineering to receive the SR-TE policies via autoroute announcement notification. These notifications are used as IGP shortcuts during SPT computation and route calculation and are installed as nexthops for applicable routes in Routing Information Base (RIB) or MPLS Forwarding Infrastructure (MFI).
<b>Cisco IOS XE Bengaluru 17.6.1</b>	
<a href="#">EVPN-IRB DHCP v4 and v6 Relay over Segment Routing</a>	<p>This feature introduces a specialised implementation of DHCP packets to support DHCPv4 and DHCPv6 in an EVPN Fabric with Distributed Anycast Gateways (DAGs) on the same Virtual Routing and Forwarding (VRF). It also avoids DHCP discovery packet floods across the fabric.</p> <p>The flooding suppression feature is also enhanced to intercept multicast or broadcast DHCP packets when DHCP relay is configured on the DAG to perform the required action and localize the scope of the service.</p> <p>This feature is not supported with Cisco ASR RSP3 module. It is only supported with Cisco ASR RSP2 module.</p> <p>This feature is only supported on NCS 4206 and NCS 4201/4202 routers.</p>
<a href="#">IS-IS Flexible Algorithm Include Affinity Support</a>	This feature supports "include-any" and "include-all" affinities in IS-IS. Prior to Cisco IOS XE Bengaluru 17.6.1 release, only Flexible Algorithm affinity "exclude-any" was supported.
<a href="#">OSPF Flexible Algorithm (Ph2): Topology-Independent Loop-Free Alternate (TI-LFA) Path</a>	This feature allows you to configure the Loop-Free Alternate (LFA) and TI-LFA backup or repair paths for a Flexible Algorithm. The backup path is computed based on the constraints and metrics of the primary path. Prior to Cisco IOS XE Bengaluru 17.6.1, OSPF Flexible Algorithm supported only the primary path.
<a href="#">SR-PCE: Enabling SR PM Delay or Liveness for PCE-Initiated Policies</a>	This feature enables the Path Computation Element (PCE) that can provision a Segment Routing Traffic Engineering (SR-TE) policy to mitigate link congestion. Prior to this release, you could only enable PM link and delay measurement using CLI-based policies. Starting with this release, you can also use PCE to enable PM link and delay measurement.
<a href="#">Stitching of Subnet Route from EVPN to L3VPN</a>	<p>This feature introduces the collapsed spine and border leaf node in the network topology of single homing DAGs with symmetric IRB, inter-subnet layer 3 traffic within fabric and inter-subnet layer 3 stitching through layer 3 border gateway. The hosts participating in fabric IRB are directly attached with the collapsed spine and border leaf node.</p> <p>This feature is not supported with Cisco ASR RSP3 module. It is only supported with Cisco ASR RSP2 module.</p> <p>This is only supported on NCS 4206 and NCS 4201/4202 routers.</p>
<b>Cisco IOS XE Bengaluru 17.5.1</b>	

Feature	Description
<a href="#">ECMP over SR-TE Policy</a>	This feature allows you to configure ECMP over SR-TE policies. In case of multiple paths, this feature enables mitigation of local congestion through load balancing.  This feature is supported on Cisco RSP3 module.
<a href="#">SR-PM Delay Deduction (Loopback Mode)</a>	This feature improves the SR-PM detection time as the PM probes are not punted on the remote nodes. Also, it does not require a third-party support for interoperability.
<a href="#">SR-TE PM: Liveness of SR Policy Endpoint</a>	This feature enables Performance Measurement (PM) liveness detection and delay measurement for an SR policy on all the segment lists of every candidate path that are present in the forwarding table using PM probes. Thus, you can easily monitor the traffic path and efficiently detect any drop of traffic due to cable or hardware or configuration failures.  This feature provides the following benefits: <ul style="list-style-type: none"> <li>• End-to-end liveness is verified before activating the candidate path in the forwarding table.</li> <li>• End-to-end liveness failure can trigger re-optimization to another path by deactivating the current path.</li> </ul>
<a href="#">Segment Routing Flexible Algorithm with OSPF</a>	This feature allows you to configure Segment Routing Flexible Algorithm with OSPF. Flexible Algorithm with OSPF supports metric minimization and avoidance, multi-plane, delay metric with rounding, and ODN with auto-steering.
<a href="#">Segment Routing Policy Counters</a>	This feature enables statistic counters to be displayed when traffic passes over the SR-TE tunnel.  You can use the command <b>show segment-routing traffic-eng policy name</b> <i>policy name</i> to view the counters.
<b>Cisco IOS XE Bengaluru 17.4.1</b>	
<a href="#">L2VPN over SR-TE Preferred Path</a>	This feature allows you to configure an SR policy as the preferred path for a VPWS or VPLS pseudowire. VPWS or VPLS pseudowires between same PEs can be routed over different SR policies based on the requirements. Prior to this release, you could only steer the traffic using the SR policy for routing IPv4 traffic to a destination pseudowire (over IGP or BGP-LU).
<a href="#">PCE Initiated SR Policy with OSPF Autoroute Announce</a>	This feature enables a steering mechanism in which IGPs automatically use the policy for destination's downstream of the policy end point.
<a href="#">Segment Routing Flexible Algorithm support for TI-LFA uLoop Avoidance, SID Leaking, and ODN with Auto-Steering</a>	This feature allows you to compute Loop Free Alternate (LFA) paths, TI-LFA backup paths, and Microloop Avoidance paths for a particular Flexible Algorithm using the same constraints as the calculation of the primary paths for such Flexible Algorithms, for IS-IS. See <a href="#">Calculation of Flexible Algorithm Path</a> .  Inter-area leaking of Flexible Algorithm SIDs and prefixes and selectively filtering the paths that are installed to the MFI are also supported. See <a href="#">Flexible Algorithm Prefix-SID Advertisement</a> and <a href="#">Installation of Forwarding Entries for Flexible Algorithm Paths</a> .

Feature	Description
Telemetry (Model-Based Telemetry and Event-Based Telemetry) Support for Performance Measurement	This feature enables Model-Based Telemetry (MDT) and Event-Based Telemetry (EDT) that allow the data to be directed to a configured receiver. This data can be used for analysis and troubleshooting purposes to maintain the health of the network. The <b>sr_5_label_push_enable</b> SDM template is mandatory for this feature to function.
<b>Cisco IOS XE Amsterdam 17.3.1</b>	
EVPN Single-Homing Over Segment Routing	The EVPN Single-Homing feature utilizes the BGP MPLS-based Ethernet VPN functionality as defined in RFC 7432. For EVPN Single-Homing feature, a CE device is attached to a single PE device and has an Ethernet Segment.
SR-TE Per-Flow (Class) ODN and Automated Steering (PCE Delegated)	This feature lets you steer traffic with SR-TE PFP based on the QoS markings on the packets. The traffic is then switched onto the appropriate path based on the forward classes of the packet. This feature is supported on the Cisco RSP2 and RSP3 modules.
Segment Routing Flexible Algorithm	Segment Routing Flexible Algorithm allows operators to customize IGP shortest path computation according to their own needs. Custom SR prefix-SIDs are assigned to forward the packets beyond link-cost-based SPF. As a result, a traffic engineered path is automatically computed by the IGP to any destination reachable by the IGP.
Segment Routing Performance Measurement Delay Measurement Using RFC 5357 (TWAMP Light)	This feature enables hardware timestamping. The Performance Measurement (PM) for link delay uses the light version of Two-Way Active Measurement Protocol (TWAMP) over IP and UDP defined in Appendix I of RFC 5357. TWAMP provides an alternative for interoperability when RFC 6374 is not used.
Segment Routing Performance Measurement End-to-End Delay Measurement	This feature allows to monitor the end-to-end delay experienced by the traffic sent over a Segment Routing policy. This feature ensures the delay does not exceed the specified threshold value and violate the SLAs. Use this feature to apply extended TE link delay metric (minimum delay value) to compute paths for Segment Routing policies as an optimization metric or as an accumulated delay bound.
Static Route Traffic Steering Using SR-TE Policy	This feature allows the non colored (BGP Extended Community) prefix to steer traffic over static policy. Prior to this release, only colored (BGP Extended Community) prefix could automatically steer traffic based on the defined policy using a tunnel interface. Unlike non colored prefix, this was possible only for the colored prefix as it could match the SR policy. IPv4 static routes are now enhanced to leverage the SR policies to aid Segment Routing Traffic Engineering (SR-TE). This facilitates traffic steering for non colored prefix as you can now configure IP Static Route with SR static policy. The following new keyword for the <b>ip route</b> command is introduced: <b>segment-routing policy</b> [ <i>policy name</i> ].

Feature	Description
<a href="#">Telemetry (Model-Based Telemetry and Event-Based Telemetry) Support for Performance Measurement</a>	This feature enables Model-Based Telemetry (MDT) and Event-Based Telemetry (EDT) that allow the data to be directed to a configured receiver. This data can be used for analysis and troubleshooting purposes to maintain the health of the network. This feature is supported on Cisco ASR 900 RSP3 module. The <b>sr_5_label_push_enable</b> SDM template is mandatory for this feature to function.
<b>Cisco IOS XE Amsterdam 17.1.1</b>	
<a href="#">EVPN-VPWS Single Homing over Segment Routing</a>	EVPN-VPWS single homing is a BGP control plane solution for point-to-point services. It has the ability to forward traffic from one network to another using Ethernet Segment without MAC lookup. EVPN-VPWS single homing works on both IP and SR core. IP core is used to support BGP while the SR core is used to switch packets between the endpoints.
<a href="#">Segment Routing Low Latency Network Slice</a>	This feature allows the advertisement and reception of the extended TE link delay metrics without any additional configuration required in IS-IS, OSPF or BGP-IS. When the link delay values are configured, they are flooded in the PCE topology and when the path computation is requested, these values are used for path calculation.
<a href="#">Segment Routing Performance Measurement Link Delay Metrics</a>	Network performance metrics such as packet loss, delay, delay variation, and bandwidth utilization is a critical measure for traffic engineering (TE) in service provider networks. These metrics provide network operators with information about characteristics of their networks for performance evaluation and helps to ensure compliance with service level agreements. The service-level agreements (SLAs) of service providers depend on the ability to measure and monitor these network performance metrics.
<a href="#">SR-TE Policy for MPLS TE</a>	The Cisco ASR 900 Series routers with Cisco RSP2 module support the newer segment routed Traffic Engineering (SR-TE) policy and you can enable the policy using the segment-routing traffic-eng command.
<a href="#">SR-TE ODN Color Extended Community for Layer 3 VPN</a>	The Cisco ASR 900 Series routers with Cisco RSP2 module support the color extended community with the following feature support: <ul style="list-style-type: none"> <li>• An egress router adds the color extended community to the BGP updates that require a Traffic-Engineered path.</li> <li>• An SR-TE policy is created on the ingress router for the color-endpoint pair.</li> </ul>





## CHAPTER 2

# Segment Routing Overview

Segment routing is a method of forwarding packets on the network based on the source routing paradigm. The forward path is determined before the packet is even sent. The path is encoded in the packet, at the source as a list of segments bearing forwarding instructions. At each hop, the top segment, which references the router information base (RIB), is used to identify the next hop. Segments are stacked in order, at the top of the packet header. When the top segment contains the identity of another node, the receiving node uses ECMP to move the packet to the next hop. When the identity is that of the receiving node, the node pops the top segment and performs the task required by the next segment.

Segment routing leverages other Interior Gateway Protocols such as IS-IS, OSPF, and MPLS for efficient and flexible forwarding. Segment routing is a faster and a more efficient way of forwarding traffic in the MPLS core network.

To understand the working of segment routing, let's understand how MPLS traffic engineering works.

- [How Does Segment Routing Work?, on page 9](#)
- [General Limitations of Segment Routing, on page 16](#)
- [Configuring Segment Routing, on page 16](#)

## How Does Segment Routing Work?

A router in a Segment Routing network can select either an explicit path or a default Interior Gateway Protocol (IGP) shortest path. Segments represent subpaths that a router can combine to form a complete route to a network destination. Each segment has an identifier (Segment Identifier) that is distributed throughout the network using new IGP extensions. The extensions are equally applicable to IPv4 and IPv6 control planes. Unlike the case for traditional MPLS networks, routers in a Segment Router network do not require Label Distribution Protocol (LDP) and Resource Reservation Protocol - Traffic Engineering (RSVP-TE) to allocate or signal their segment identifiers and program their forwarding information.

Each router (node) and each link (adjacency) has an associated segment identifier (SID). **Node segment identifiers** are globally unique and represent the shortest path to a router as determined by the IGP. The network administrator allocates a node ID to each router from a reserved block. On the other hand, **adjacency segment identifiers** are locally significant and represent a specific adjacency, such as egress interface, to a neighboring router. Routers automatically generate adjacency identifiers outside of the reserved block of node IDs. In an MPLS network, a segment identifier is encoded as an MPLS label stack entry. Segment IDs direct the data along a specified path. There are two kinds of segment IDs:

- **Prefix SID**—A segment ID that contains an IP address prefix calculated by an IGP in the service provider core network. Prefix SIDs are globally unique. A node SID is a special form of prefix SID that contains

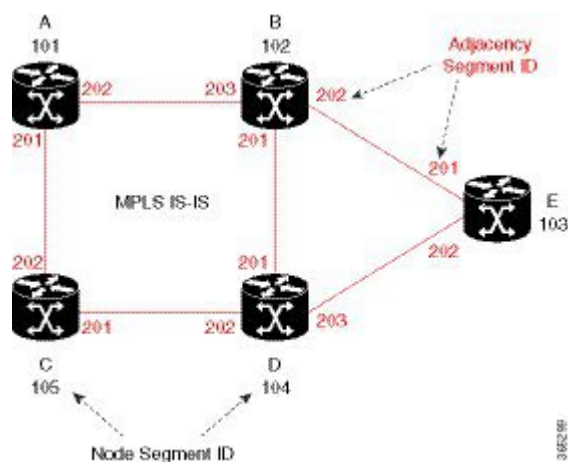
the loop-back address of the node as the prefix. It is advertised as an index into the node-specific SR Global Block or SRGB.

- **Adjacency SID**—A segment ID that contains an advertising router's adjacency to a neighbor. An adjacency SID is a link between two routers. Since the adjacency SID is relative to a specific router, it is locally unique.

## Examples for Segment Routing

The following figure illustrates an MPLS network with five routers using Segment Routing, IS-IS, a label range of 100 to 199 for node IDs, and 200 and higher for adjacency IDs. IS-IS would distribute IP prefix reachability alongside segment ID (the MPLS label) across the network.

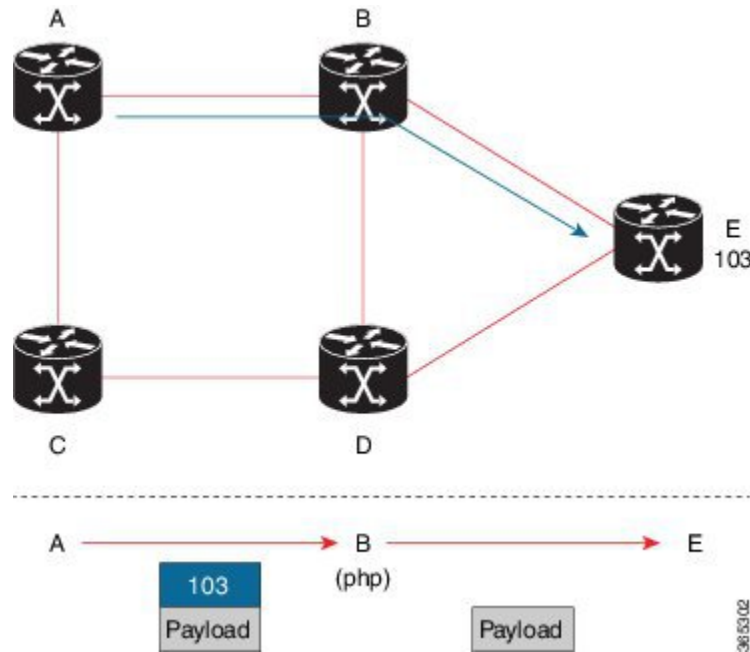
**Figure 1: An MPLS Network with Five Routers Using Segment Routing**



In the previous example, any router sending traffic to router E would push label 103 (router E node segment identifier) to forward traffic using the IS-IS shortest path. The MPLS label-swapping operation at each hop preserves label 103 until the packet arrives at E (Figure 2). On the other hand, adjacency segments behave differently. For example, if a packet arrives at Router D with a top-of-stack MPLS label of 203 (D-to-E adjacency segment identifier), Router D pops the label and forward the traffic to Router E.

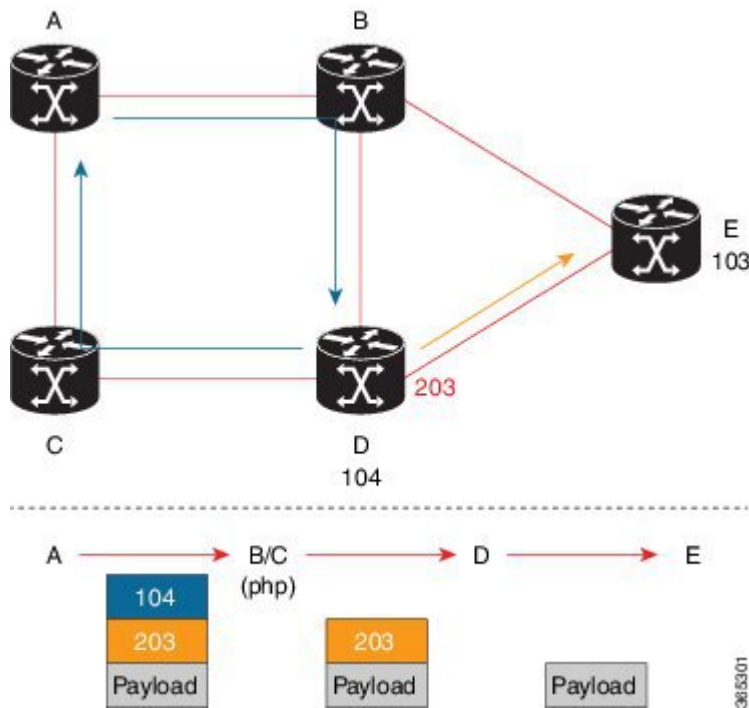


Figure 2: MPLS Label-Swapping Operation



Segment identifiers can be combined as an ordered list to perform traffic engineering. A segment list can contain several adjacency segments, several node segments, or a combination of both depending on the forwarding requirements. In the previous example, Router A can alternatively push label stack (104, 203) to reach Router E using the shortest path and all applicable ECMPs to Router D, and then through an explicit interface onto the destination (Figure 3). Router A does not need to signal the new path, and the state information remains constant in the network. Router A ultimately enforces a forwarding policy that determines which flows destined to router E are switched through a particular path.

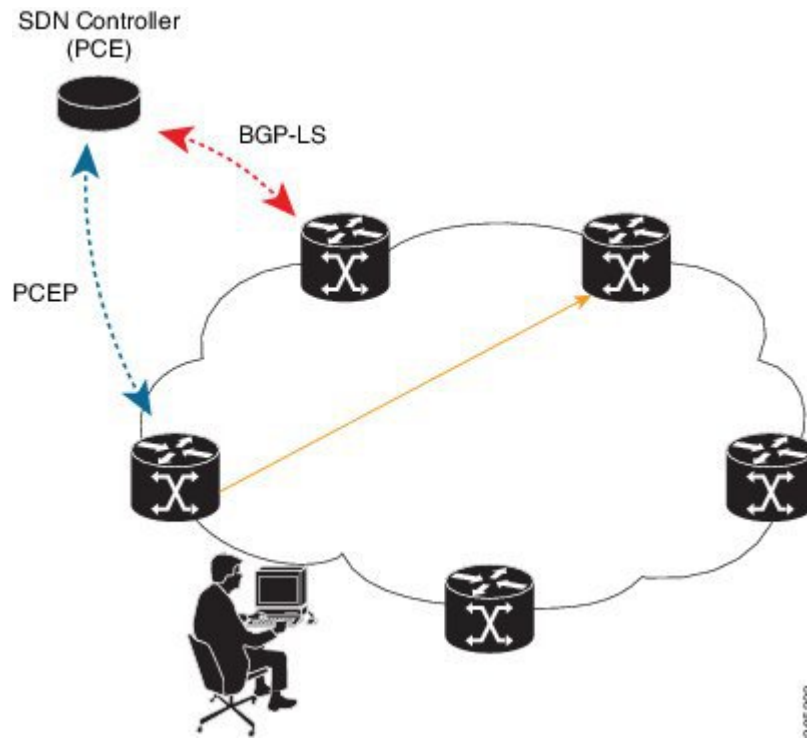
Figure 3: Destination Path of Router E



## Benefits of Segment Routing

- **Ready for SDN**—Segment Routing is a compelling architecture conceived to embrace Software-Defined Network (SDN) and is the foundation for Application Engineered Routing (AER). It strikes a balance between network-based distributed intelligence, such as automatic link and node protection, and controller-based centralized intelligence, such as traffic optimization. It can provide strict network performance guarantees, efficient use of network resources, and high scalability for application-based transactions. The network uses minimal state information to meet these requirements. Segment routing can be easily integrated with a controller-based SDN architecture. The following figure illustrates a sample SDN scenario where the controller performs centralized optimization, including bandwidth admission control. In this scenario, the controller has a complete picture of the network topology and flows. A router can request a path to a destination with certain characteristics, for example, delay, bandwidth, diversity. The controller computes an optimal path and returns the corresponding segment list, such as an MPLS label stack, to the requesting router. At that point, the router can inject traffic with the segment list without any additional signaling in the network.

Figure 4: SDN Controller

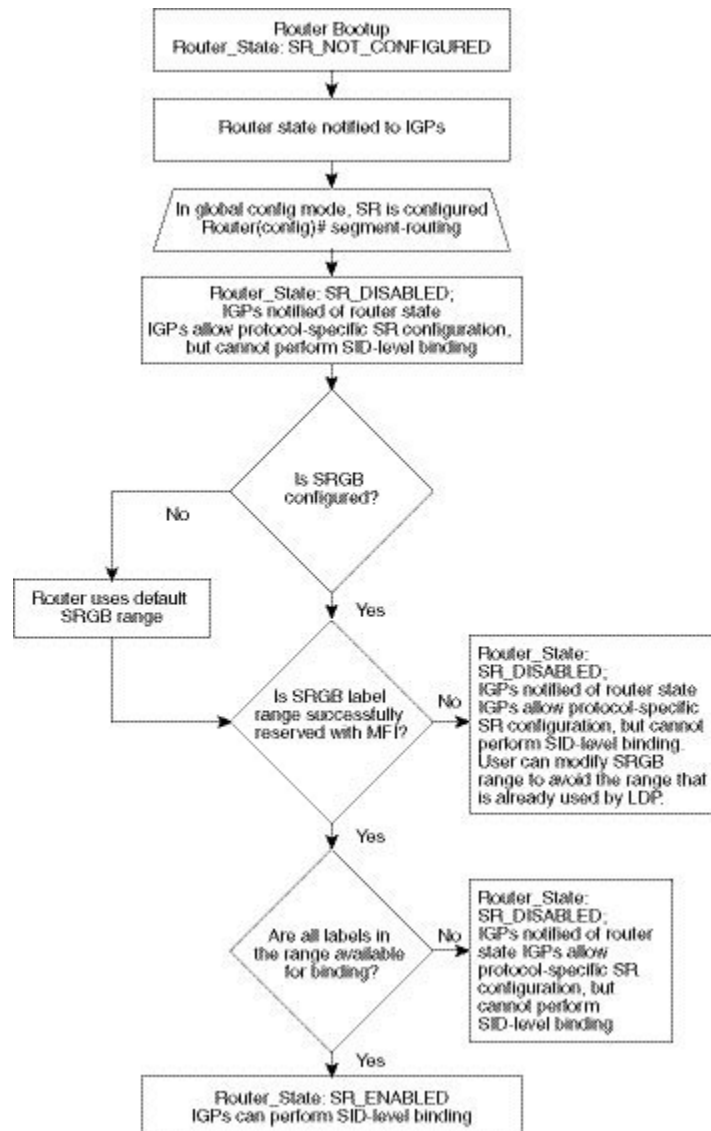


- In addition, segment lists allow complete network virtualization without adding any application state to the network. The state is encoded in the packet as a list of segments. Because the network only maintains segment state, it can support a large number - and a higher frequency - of transaction-based application requests without creating any burden on the network.
- Simplified—
  - When applied to the MPLS data plane, Segment Routing offers the ability to tunnel MPLS services (VPN, VPLS, and VPWS) from an ingress provider edge to an egress provider edge without any other protocol than an IGP (IS-IS or OSPF).
  - Simpler operation without separate protocols for label distribution (for example, no LDP or RSVP).
  - No complex LDP or IGP synchronization to troubleshoot.
  - Better utilization of installed infrastructure, for lower capital expenditures (CapEx), with ECMP-aware shortest path forwarding (using node segment IDs).
- Supports Fast Reroute (FRR)—Deliver automated FRR for any topology. If the link or node fails in a network, MPLS uses the FRR mechanism for convergence. With segment routing, the convergence time is sub-50-msec.
- Large-scale Data Center—
  - Segment Routing simplifies MPLS-enabled data center designs using Border Gateway Protocol (BGP) RFC 3107 - IPv4 labeled unicast among Top-of-the-Rack/Leaf/Spine switches.
  - BGP distributes the node segment ID, equivalent to IGP node SID.

- Any node within the topology allocates the same BGP segment for the same switch.
- The same benefits are provided as for IGP node SID: ECMP and automated FRR (BGP PIC (Prefix Independent Convergence)).
- This is a building block for traffic engineering - SR TE data center fabric optimization.
- Scalable—
  - Avoid thousands of labels in LDP database.
  - Avoid thousands of MPLS Traffic Engineering LSPs in the network.
  - Avoid thousands of tunnels to configure.
- Dual-plane Networks—
  - Segment Routing provides a simple solution for disjointness enforcement within a so-called “dual-plane” network, where the route to an edge destination from a given plane stays within the plane unless the plane is partitioned.
  - An additional SID “anycast” segment ID allows the expression of macro policies such as: “Flow 1 injected in node A toward node Z must go through plane 1” and “Flow 2 injected in node A toward node Z must go through plane 2.”
- Centralized Traffic Engineering—
  - Controllers and orchestration platforms can interact with Segment Routing traffic engineering for centralized optimization, such as WAN optimization.
  - Network changes such as congestion can trigger an application to optimize (recompute) the placement of segment routing traffic engineering tunnels.
  - Segment Routing tunnels are dynamically programmed onto the network from an orchestrator using southbound protocols like PCE.
  - Agile network programming is possible since Segment Routing tunnels do not require signaling and per-flow state at midpoints and tailend routers.
- Egress Peering Traffic Engineering (EPE)—
  - Segment Routing allows centralized EPE.
  - A controller instructs an ingress provider edge and content source to use a specific egress provider edge and specific external interface to reach a destination.
  - BGP “peering” segment IDs are used to express source-routed inter-domain paths.
  - Controllers learn BGP peering SIDs and the external topology of the egress border router through BGP Link Status (BGP-LS) EPE routes.
  - Controllers program ingress points with a desired path.
- Plug-and-play deployment—Segment routing tunnels are inter-operable with existing MPLS control and data planes and can be implemented in an existing deployment.

## Segment Routing Limitations

Figure 5: Segment Routing State Flow



- Segment Routing must be globally enabled on the chassis *before* enabling it on the IGP's, like IS-IS or OSPF.
- Segment routing must be configured on the IS-IS instance *before* configuring a prefix SID value.
- The prefix SID value must be removed from all the interfaces under the same IS-IS instance before disabling segment routing.

## General Limitations of Segment Routing

- The following types of services are supported:
  - VPLS
  - Layer 2 VPN
  - Layer 3 VPN
  - Global Prefixes
- To cater to scaled services, ensure that you use the following values:
  - 1500 IGP
  - 4000 L3VPN
  - 2000 L2VPN virtual circuits
- Ensure that you have the micro-loop avoidance rib-update-delay under IGP configuration as 20,000 msec. The nodes on the post-convergence path may need this time to update the forwarding plane. The number of affected prefixes may also cause a delay.
- While configuring FRR, *carrier-delay down msec 0* must be configured under interface configuration.
- Cisco ASR 900 Router with RSP3 module can push a maximum of 4 MPLS labels in the egress direction. This includes service labels (L3VPN, L2VPN, 6PE, 6VPE), RFC 3107 BGP-LU label and SR labels for FRR primary/backup paths. This restriction is not applicable while using Segment Routing Traffic Engineering. For information see [Segment Routing Traffic Engineering, on page 51](#).

## Configuring Segment Routing

To configure segment routing

1. Globally enable segment routing:

```
enable
configure terminal
segment-routing mpls
```

2. Specify the range of MPLS labels to be used to instantiate the segment routing SIDs into MPLS data plane.

```
global block
16-8000
```

3. Associate SID values with local prefix values.

```
connected-prefix-sid-map
address-family ipv4
  92.0.0.0/24 index 51 range 1
  2.2.2.2/32 index 2 range 1
```

## Configuring Segment Routing on an IGP Instance

```
enable
configure terminal
segment-routing
segment-routing [area N] {mpls | disable}
```

This command enables MPLS on all interfaces and programs the MPLS labels for forwarding.



---

**Note** If the **area** keyword is specified, segment routing is enabled only on that area.

---



---

**Note** The **disable** keyword can be used only if the **area** keyword is specified.

---

## Enabling Advertisement of Mapping Server Prefix Ranges

Global segment routing configuration may contain prefix-to-SID mapping entries for prefixes that are not local to the router. Each of these entries specifies a range of prefixes. Remote mapping entries can be used to find SIDs for prefixes connected to routers that do not support SR and hence not capable of advertising SIDs themselves. This capability is part of SR-LDP inter-working functionality. OSPF learns the ranges configured in the global SR configuration and advertises them in the Extended Prefix Range TLVs.

To permit an OSPF instance to advertise mapping entries configured in the global SR mode, use the following command in router mode.



---

**Note** By default, this command is disabled. That is, no mapping ranges are advertised by OSPF even if they are configured in the global SR mode.

---

```
segment-routing prefix-sid-map advertise-local
```

## Disabling the Mapping Server

When computing SIDs for prefixes, IGP considers the prefix ranges received from mapping servers in the network by default. However, if this functionality needs to be disabled, use the following command in router mode. If the SR-LDP feature needs to be disabled, it is done in router-mode:

```
no segment-routing prefix-sid-map receive
```



---

**Note** This command does not affect processing of 'native' SIDs, that is, those SIDs that are advertised in the Extended Prefix TLVs by routers to whom the prefix is locally connected.

---







## CHAPTER 3

# Segment Routing Global Block (SRGB)

The Segment Routing Global Block (SRGB) is the range of label values reserved for Segment Routing (SR) in the Label Switching Database (LSD). These values are assigned as segment identifiers (SIDs) to SR-enabled nodes and have global significance throughout the domain. On SR-capable routers, SRGB is enabled by default so that label values are automatically reserved when the router first boots whether SR is enabled or not. The default SRGB range is 16000 to 23999, and the SRGB can be disabled if SR is not used. You can also specify separate SRGBs for IS-IS and OSPF protocols so long as the ranges do not overlap.

- [SRGB Limitations, on page 19](#)
- [Configuring the SRGB, on page 19](#)
- [Adjacency Segment Identifiers, on page 20](#)
- [Prefix Segment Identifiers, on page 20](#)

## SRGB Limitations

- LSD label values 0-15,999 are reserved.
- The SRGB size cannot be more than  $2^{16}$  or 65536.
- The SRGB upper bound value cannot exceed the platform capability.
- The SRGB cannot be configured to be the same value as the default SRGB. So SRGB cannot be configured for 16000 to 23999.



**Note** Label values that are not previously reserved are available for dynamic assignment.

## Configuring the SRGB

When SR is enabled, but the SRGB is not configured, the system reserves a default label range. After you configure the SRGB, the default label range is released.

When the SR is disabled, the corresponding label range is released and a default label range is reserved.

## Reserving a New SRGB

If the SRGB is reserved successfully with the MPLS Forwarding Interface (MFI), then SR is enabled. Otherwise, SR is disabled, but the segment routing MPLS global configuration is saved.

If the MFI reserves the SRGB, but there are some dynamically allocated labels in use in that range, then SR is disabled, the SRGB is released, and the configuration is saved.

## Modifying an Existing SRGB

There can be two scenarios when modifying an SRGB:

1. Extending the upper bound of an existing SRGB. For example,

Current range—16000 to 20000

New range—16000 to 24000

The new SRGB range is reserved based on the conditions mentioned in [Reserving a New SRGB](#).



**Note** In all cases, the new configuration is always saved.

2. The new range is either a disjoint range from the existing SRGB or overlaps with existing SRGB

In this case, SR is disabled and the previous configuration is deleted. See the figure in [Segment Routing Limitations, on page 15](#) section.

To configure an SRGB:

```
enable
configure terminal
segment-routing mpls
global-block 16000 17000
```

## Adjacency Segment Identifiers

The Adjacency Segment Identifier (adj-SID) is a local label that points to a specific interface and a next hop out of that interface. The router in which the interface is installed automatically assigns an SR SID to the connection using values outside the range of those in its segment routing global block. This SID has local significance only and is not distributed for inclusion in the FIBs of other routers.



**Note** Only IPv4 address-family supports allocating adjacency-SIDs.

## Prefix Segment Identifiers

A prefix segment identifier (SID) identifies a segment routing tunnel leading to the destination represented by a prefix. The maximum prefix SID value is  $(2^{16}) - 1$ .

When a prefix SID value  $x$  is configured, the prefix SID translates to a label value equivalent to  $x + \text{lower boundary of SRGB}$ . For example, if the default SRGB is used, configuring a prefix-SID of 10 for interface Loopback 0 with IPv4 address 1.0.0.1/32 results in assigning the label 16010 to the prefix 1.0.0.1/32.

## Types of Prefix SID Mapping

- Export Prefix SID Mapping
- Connected Prefix SID Mapping

### Export Prefix SID Mapping

These are the prefix-sid mappings that the IGP exports to the segment routing nodes through the SID or the Label Binding TLVs, and define the prefix-SID bindings for the network.

### Connected Prefix SID Mappings

These are the prefix SID mappings that define the Prefix SID or Node-SID of the interface. Connected prefix SID mappings are created under the mapping-server that the IGPs query to obtain the prefix SID of the interface of an SR node.

## Configuring Export Prefix SID Mapping

```
segment-routing mpls
connected-prefix-sid-map
address-family ipv4 cisco
  <prefix> / <masklen> [index | absolute] <label> [range <value>]
```

For example,

```
connected-prefix-sid-map
address-family ipv4
92.0.0.0/24 index 51 range 1
2.2.2.2/32 index 2 range 1
```



**Note** The **index** and **absolute** keywords indicate whether the label value entered should be interpreted as an index in the SRGB or as an absolute value. By default, the value is interpreted as an index value. Regardless of the configuration mode, the value is advertised by the protocols as an index value.

## Configuring Connected Prefix SID Mapping

```
segment-routing mpls
mapping-server
  prefix-sid-map
    address-family ipv4
cisco
  <prefix> / <masklen> [index
```

```
| absolute
] <label> [range <value>]
```

**Note**

The **index** and **absolute** keywords indicate whether the label value entered should be interpreted as an index in the SRGB or as an absolute value. By default, the value is interpreted as an index value. Regardless of the configuration mode, the value is advertised by the protocols as an index value.

## Configuring the Preferred Label

The Cisco Express Forwarding engine (CEF) can use either the prefix labels stored in the routing information base (RIB) or the labels provided by the label distribution protocol (LDP), for forwarding.

You can set the prefix label or the LDP label as the preferred label by using:

```
segment-routing mpls
  set-attributes
    address-family ipv4
cisco
  sr-label-preferred
```

The IGP communicates the preference to the CEF, which in turn processes the forwarding for that prefix accordingly.

## Configuring the Segment Routing Explicit Null

To ensure that quality of service is maintained between the penultimate and last-hop routers, the penultimate router swaps the segment routing label with an explicit null label and forwards the packet to the last-hop router. This swapping preserves the MPLS header attributes for QoS purposes.

```
segment-routing mpls
set-attributes
address-family ipv4 cisco
explicit-null
```



## CHAPTER 4

# Using Segment Routing with IS-IS

We know that segment routing enables a node to select any path (explicit or derived from the computations of the internal gateway protocol's shortest path). This path is not dependent on a hop-by-hop signaling technique (through LDP or RSVP), but on a set of segments that are advertised by a routing protocol, such as IS-IS or OSPF. These segments act as topological sub-paths that can be combined to form the desired path.

Segment Routing must be enabled before any IGP, such as IS-IS or OSPF, can configure segment routing functionality. Similarly, when segment routing is disabled, all IGP-related configuration is also disabled.

- [Restrictions for Using Segment Routing with IS-IS, on page 23](#)
- [Enabling Segment Routing, on page 23](#)
- [Enabling Segment Routing for IGPs, on page 24](#)
- [Prefix-SID Received in LSPs from Remote routers, on page 25](#)
- [Segment Routing Adjacency SID Advertisement, on page 25](#)
- [Segment Routing Mapping Server \(SRMS\), on page 26](#)
- [SRGB Range Changes, on page 27](#)
- [MPLS Forwarding on an Interface, on page 28](#)
- [Segment Routing and LDP Preference, on page 28](#)
- [Segment Routing-TE, on page 28](#)
- [RLFA LDP and SR, on page 28](#)
- [Topology-Independent LFA, on page 29](#)

## Restrictions for Using Segment Routing with IS-IS

- Effective Cisco IOS XE Release 3.16S, ISIS supports segment routing for IPv4 only.
- Segment routing must be configured at the top level before any routing protocol configuration is allowed under its router configuration sub mode.
- IS-IS protocol SR command is based on per topology (IPv4 address family).
- Only network type = point-to-point is supported.

## Enabling Segment Routing

There are two levels of configuration required to enable segment routing for a routing protocol instance. The top level segment routing configuration which is managed by segment routing infrastructure component

enables segment routing, whereas, segment routing configuration at the router level enables segment routing for a specific address-family of a routing protocol instance.

There are three segment routing states:

- SR\_NOT\_CONFIGURED
- SR\_DISABLED
- SR\_ENABLED

Segment routing configuration under the IGPs is allowed only if the SR state is either SR\_DISABLED or SR\_ENABLED. The SR\_ENABLED state indicates that there is at least a valid SRGB range reserved through the MFI successfully.

## Enabling Segment Routing for IGPs

You can enable segment routing for IGPs under the router configuration sub mode, through commands. However, IGP segment routing are enabled only after the global SR is configured.



### Note

IS-IS protocol SR command is based on per topology (IPv4 address family).

The SR\_ENABLED is a necessary state for any protocol to enable SR, however, it is not a sufficient for enabling SR for a protocol instance. The reason being that the IS-IS still does not have any information about segment routing global block (SRGB) information. When the request to receive information about the SRGB is processed successfully, the IS-IS SR operational state is enabled

Segment Routing requires each router to advertise its segment routing data-plane capability and the range of MPLS label values that are used for segment routing in the case where global SIDs are allocated.

Data-plane capabilities and label ranges are advertised using the SR-capabilities sub-TLV inserted into the IS-IS Router Capability TLV-242 that is defined in RFC4971.

ISIS SR-capabilities sub TLV includes all reserved SRGB ranges. However, the Cisco implementation supports only one SRGB range.

The supported IPv4 prefix-SID sub TLV are TLV-135 and TLV-235.

## Configuring Segment Routing on IS-IS

This section describes configuring segment routing IPV4 for IS-IS protocol under the router configuration sub mode.

```
[no] segment-routing mpls
```



### Note

This command is allowed only when segment routing is configured at the top level.

The following is an example of configuring IS-IS segment routing:

```
segment-routing mpls
```

```
router isis
net 33.0001.0001.0001.00
metric-style wide
segment-routing mpls
passive-interface Loopback2
```

## Prefix-SID Received in LSPs from Remote routers

Prefix SIDs received in a label switched path (LSP) with a reachability TLV (TLV 135 and 235) are downloaded to the routing information base (RIB) if all of the following conditions are met:

- Segment routing is enabled for the topology and address-family
- Prefix-SID is valid
- The local label binding to MFI is successful.

For a prefix-SID received with reachability TLVs (TLV 135 and 235), the label is downloaded through RIB the same way as BGP downloads per prefix VPN labels.

If the path is a remote LFA path, ISIS downloads the path the same way it downloads it before adding the segment routing functionality but does *not* download any label with this path. This behavior ensures that remote LFA functionality is still supported using LDP.

## Limitations

- For SIDs that do not fit in the specified SID range, labels are not used when updating the RIB. For cases, where SID does fit in the SID range, but does not fit the next-hop neighbor SID range, remote label associated with that path is not installed.
- Node SIDs received in an LSP with reachability TLVs (TLV 135 and 235) are downloaded to RIB only if segment routing is enabled under the corresponding address-family.
- In case of multiple best next hops, if not all next hops support segment routing, ISIS will treat this instance similar to when mismatched labels are assigned to the same prefix. That is, IS-IS ignores the labels and installs unlabeled paths for all ECMP paths into the global RIB.

## Segment Routing Adjacency SID Advertisement

Effective with Cisco IOS XE Release 3.17S, IS-IS supports the advertisement of segment routing Adjacency SID. An Adjacency Segment Identifier (Adj-SID) represents a router adjacency in Segment Routing.

A segment routing-capable router may allocate an Adj-SID for each of its adjacencies and an Adj-SID sub-TLV is defined to carry this SID in the Adjacency TLVs. IS-IS adjacencies are advertised using one of the IS-Neighbor TLVs below:

- TLV-22 [RFC5305]
- TLV-23 [RFC5311]

IS-IS allocates the adjacency SID for each IS-IS neighbor only if the IS-IS adjacency state is up and IS-IS segment routing internal operational state is enabled. If an adjacency SID allocation failure is due to out-of-label resource, IS-IS retries to allocate the Adj-SID periodically in a default interval (30 seconds).

## Multiple Adjacency-SIDs

Effective with Cisco IOS XE Release 3.18S, multiple adjacency-SIDs are supported. For each protected P2P/LAN adjacency, IS-IS allocates two Adj-SIDs. The backup Adj-SID is only allocated and advertised when FRR (local LFA) is enabled on the interface. If FRR is disabled, then the backup adjacency-SID is released. The persistence of protected adj-SID in forwarding plane is supported. When the primary link is down, IS-IS delays the release of its backup Adj-SID until the delay timer expires. This allows the forwarding plane to continue to forward the traffic through the backup path until the router is converged.

Cisco IOS XE Release 3.18S, IS-IS Adj-SID is changed to be per level based since the forwarding plane is unaware of protocol-specific levels. The allocated and advertised backup Adj-SIDs can be displayed in the output of **show isis neighbor detail** and **show isis data verbose** commands.

## Segment Routing Mapping Server (SRMS)

Segment Routing Mapping Server (SRMS) allows configuration and maintenance of the Prefix-SID mapping policy entries. Effective with Cisco IOS XE Release 3.17S, the IGP uses the active policy of the SRMS to determine the SID values when programming the forwarding plane.

The SRMS provides prefixes to SID/Label mapping policy for the network. IGPs, on the other hand, are responsible for advertising prefixes to SID/Label mapping policy through the Prefix-SID/Label Binding TLV. Active policy information and changes are notified to the IGPs, which use active policy information to update forwarding information.

## Connected Prefix SIDs

Sometimes, a router may install a prefix with a SID that is different than what it advertises to the LSP. For example, if more than one protocol or more than one IGP instance is announcing the same prefix with different SIDs to the SRMS, the SRMS resolves the conflict and announces the winning prefix and SID that may not be the same as the local instance. In that case, the IGP always advertises what it learns from its source LSP although it still tries to install the SID which may be different than what it learns in its LSP. This is done to prevent the IGP from redistributing the SIDs from another protocol or another protocol instance.

## Configuring IS-IS SRMS

The following command enables the IS-IS SRMS and allows IS-IS to advertise local mapping entries. IS-IS does not send remote entries to the SRMS library. However, IS-IS uses the SRMS active policy, which is computed based only on the locally configured mapping entries.

```
[no] segment-routing prefix-sid-map advertise-local
```

## Configuring IS-IS SRMS Client

By default, the IS-IS SRMS client mode is enabled. IS-IS always sends remote prefix-sid-mapping entries received through LSP, to SRMS. The SRMS active policy is calculated based on both, local and remote mapping entries.

The following command disables the prefix-sid-mapping client functionality.



```
segment-routing prefix-sid-map receive [disable]
```

This command is configured on the receiver side.

## Configuring ISIS SID Binding TLV Domain Flooding

By default, the IS-IS SRMS server does not flood SID binding entries within the routing domain. In Cisco IOS XE Release 3.18S, the optional keyword **domain-wide** in the IS-IS SRMS server mode command to enable the SID and Label binding TLV flooding functionality.

```
segment-routing prefix-sid-map advertise-local [domain-wide]
```

The **domain-wide** keyword enables the IS-IS SRMS server to advertise SID binding TLV across the entire routing domain.



---

**Note** The option is valid only if IS-IS SRMS performs in the SRMS server mode.

---

## SRGB Range Changes

When IS-IS segment routing is configured, IS-IS must request an interaction with the SRGB before IS-IS SR operational state can be enabled. If no SRGB range is created, IS-IS will not be enabled.

When an SRGB change event occurs, IS-IS makes the corresponding changes in its sub-block entries. IS-IS also advertises the newly created or extended SRGB range in SR-capabilities sub-TLV and updates the prefix-sid sub TLV advertisement.



---

**Note** In Cisco IOS XE Release 3.16S only one SRGB range and SRGB extension for the modification are supported.

---

## SRGB Deletion

When IS-IS receives an SRGB deletion event, it looks for an SRGB entry in the IS-IS SRGB queue list. If an SRGB entry does not exist, IS-IS makes sure that there is no pending SRGB created event. If a pending SRGB creation event is found, then IS-IS removes the SRGB creation event, and completes the SRGB delete processing,

If an SRGB entry is found in the IS-IS SRGB queue, IS-IS locks the SRGB, redistributes the RIBs and un-advertises all prefixed-SIDs that have SID value within the pending delete SRGB range, and un-advertises the SRGB range from SR-capabilities sub TLV. Once IS-IS has completed the SRGB deletion processing, it unlocks the SRGB and deletes the SRGB from its SR sub-block entry.

If there is no valid SRGB after the deletion of the SRGB, IS-IS SR operational state becomes disabled.

## MPLS Forwarding on an Interface

MPLS forwarding must be enabled before segment routing can use an interface. IS-IS is responsible for enabling MPLS forwarding on an interface.

When segment routing is enabled for a IS-IS topology, or IS-IS segment routing operational state is enabled, IS-IS enables MPLS for any interface on which the IS-IS topology is active. Similarly, when segment routing is disabled for a IS-IS topology, IS-IS disables the MPLS forwarding on all interfaces for that topology.

## Segment Routing and LDP Preference

In Cisco IOS XE Release 3.16S, the command **sr-prefer** is used to tell the forwarding interface to prefer using segment routing labels over LDP labels for all prefixes in a topology.

## Segment Routing-TE

Segment Routing Traffic Engineering requires the IGP to provide segment routing related information to TE. The information includes SRGB, Adjacency-SID, Prefix-SID, primary and repair paths for all nodes in the topology.

The maximum number of allowed SR-TE tunnels are 510.

## Enabling and Disabling SR-TE Announcements

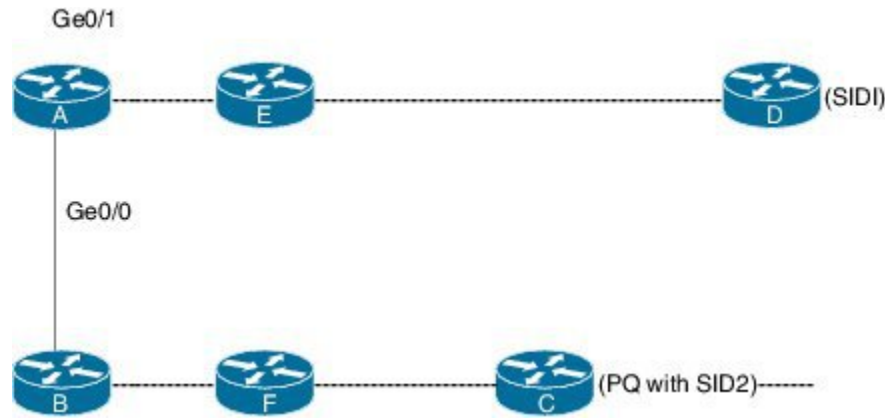
IS-IS announces the SR information to TE when it detects that both, IS-IS SR and TE are enabled for at least one level. IS-IS announce only the information that is obtained from the level for which TE is configured.

Similarly, IS-IS instructs TE to delete all announcements when it detects that SR is not enabled or TE is no longer configured on any level.

## RLFA LDP and SR

Consider the following topology.

Figure 6: Sample Topology



The traffic flows from A to D. The primary path is A-E-D and the primary next hop interface is Ge0/1. The secondary path is A-B-F-C-D, and C is the PQ node. The repair tunnel ends at PQ node C. The existing RLFA uses LDP TE tunnel for the repair path. When both LDP and SR are enabled, the LDP tunnel is used for RLFA repair path by default unless the segment routing preferred is configured through the **sr-prefer** command.

## Topology-Independent LFA

When the local LFA and remote LFA are enabled, there is a good coverage of the prefixes to be protected. However, for some rare topologies that do not have a PQ intersect node, both local and remote LFA will fail to find a release node to protect the failed link. Furthermore, there is no way to prefer a post-convergence path, as the two algorithms have no knowledge of the post-convergence characteristics of the LFA.

To overcome the above limitation, effective Cisco IOS XE Release 3.18S, topology-independent LFA (TI-LFA) is supported on an SR-enabled network.

In Cisco IOS XE Release 3.18S, TI LFA supports the following:

- Link Protection—The LFA provides repair path for failure of the link.
- Local LFA—Whenever a local LFA on the post convergence path is available, it is preferred over TI-LFA because local LFA does not require additional SID for the repair path. That is, the label for the PQ node is not needed for the release node.
- Local LFA for extended P space—For nodes in the extended P space, local LFA is still the most economical method for the repair path. In this case, TI-LFA will not be chosen.
- Tunnel to PQ intersect node—This is similar to remote LFA except that the repair path is guaranteed on the post convergence path using TI-LFA.
- Tunnel to PQ disjoint node—This capability is unique to the TI-LFA in the case when local and remote LFA cannot find a repair path.
- Tunnel to traverse multiple intersect or disjoint PQ nodes, up to the platform's maximum supported labels—TI-LFA provides complete coverage of all prefixes.
- P2P interfaces for the protected link—TI-LFA protects P2P interfaces.
- Asymmetrical links—The ISIS metrics between the neighbors are not the same.
- Multi-homed (anycast) prefix protection—The same prefix may be originated by multiple nodes.
- Protected prefix filtering—The route-map includes or excludes a list of prefixes to be protected and the option to limit the maximum repair distance to the release node.
- Tiebreakers—A subset of existing tiebreakers, applicable to TI-LFA, is supported.

## Restrictions for the TI-LFA

- IGP throttles timers that are required for RLFA tunnel are also applicable to SR and SR TI-LFA.
- BFD is not supported with SR and TI-LFA.
- Scale values supported for TI-LFA
  - Global prefixes: 1500
  - L3VPN: 4000 prefixes
  - L2VPN: 2000 virtual circuits
- SR and TI-LFA are supported on BDI and routed ports.
- Four MPLS label push is supported. TI-LFA tunnel carries a maximum of two labels and the other two labels are for services.
- RSP3 module support a maximum of 2 labels under TI-LFA tunnel.

## Tie-breaker

Local and remote LFA use default or user-configured heuristics to break the tie when there is more than one path to protect the prefix. The attributes are used to trim down the number of repair paths at the end of the TI-LFA link protection computation before the load balancing.

Local LFA and remote-LFA support the following tiebreakers:

- linecard-disjoint—Prefers the line card disjoint repair path
- lowest-backup-path-metric—Prefers the repair path with lowest total metric
- node-protecting—Prefers node protecting repair path
- srlg-disjoint—Prefers SRLG disjoint repair path
- load-sharing—Distributes repair paths equally among links and prefixes

For TI-LFA link protection, the following tiebreakers are supported:

- linecard-disjoint—Prefers the line card disjoint repair path.

How it works: When there are two repair paths for a particular prefix, the path that the output port on different line card than that of the primary port is chosen as the repair path.

The following variant of the linecard-disjoint is supported:

- LC disjoint index—If both the repair paths are on the same line card as that of the primary path, then, both paths are considered as candidates. If one of the path is on a different line card, then that path is chosen as the repair path.
- srlg-disjoint—Prefers the SRLG disjoint repair path

The SRLG ID can be configured for each interface. When there are two repair paths for a prefix, the configured SRLG ID for the repair path is compared with that of the primary path SRLG ID. If the SRLG IDs for the secondary path is different than that of the primary, that path is chosen as the repair path.



**Note** This policy comes into effect only when the primary path is configured with an SRLG ID.

The following variant of the `srlg-disjoint` is supported:

- `srlg index`—If both the repair paths have the same SRLG ID as that of the primary path, then, both the paths are considered as candidates. If one of the path has a different srlg id, then path is chosen as the repair path.
- `node-protecting`—For TI-LFA node protection, the protected node is removed when computing the post-convergence shortest path. The repair path must direct traffic around the protected node.

It is possible to configure both node and SRLG protection modes for the same interface or the same protocol instance. In that case, an additional TI-LFA node-SRLG combination protection algorithm is run. The TI-LFA node-SRLG combination algorithm removes the protected node and all members of the interface with the same SRLG group when computing the post-convergence SPT.

For TI-LFA node protection, SRLG protection, and node-SRLG combination protection, it is likely the coverage for the protected prefixes is small. TI-LFA link protection is also run to provide coverage for the prefixes that not yet covered. However, optimization can be achieved when SRLG protection is enabled with no SRLG group on the interface. In that case, SRLG protection produces the same result as link protection and link protection is skipped. Furthermore, if node-protection is also configured in this case, TI-LFA node-SRLG combination protection produces the same result as node-protection and node-protection is skipped.

## Interface FRR Tiebreakers

For TI-LFA node and SRLG protection, interface FRR tiebreakers must also be provided. Existing FRR tiebreakers are configured on a per protocol instance. Because FRR tiebreakers are not specific to TI-LFA, interface FRR tiebreakers are available for all FRR types. When both interface and protocol instance FRR tiebreakers are configured, the interface FRR tiebreakers take precedence over the protocol instance. When interface FRR tiebreakers are not configured, the interface inherits the protocol instance FRR tiebreakers. As with the existing tiebreakers, the priority must be unique among the interface and protocol instance for the tiebreakers.

The following interface FRR tiebreaker commands apply only to the particular interface.

```
isis fast-reroute tie-break
[level-1 | level-2] linecard-disjoint
priority
isis fast-reroute tie-break
[level-1 | level-2] lowest-backup-metric
priority
isis fast-reroute tie-break
[level-1 | level-2] node-protecting
priority
isis fast-reroute tie-break
[level-1 | level-2] srlg-disjoint
priority
isis fast-reroute tie-break
[level-1 | level-2] default
```

Tie-breaker default and explicit tie-breaker on the same interface are mutually exclusive.

The following tie-breakers are enabled by default on all LFAs:

- linecard-disjoint
- lowest-backup-metric
- srlg-disjoint

Effective with Cisco IOS XE Release 3.18S, node-protecting tie-breaker is disabled by default.

## Limitations on Tie-Breakers

The following tie-breakers are not applicable for these LFA scheme.

TILFA:

- broadcast-interface-disjoint
- downstream
- primary-path
- secondary-path

RLFA:

- broadcast-interface-disjoint
- node-protecting
- downstream
- primary-path
- secondary-path

## Configuring T1 LFA

TI-LFA is disabled by default. There are two methods to enable TI-LFA:

1. Using protocol enablement—Enable TI-LFA in router isis mode. This enables TI-LFA for all ISIS interfaces. Optionally, use the interface command to exclude the interfaces on which TI-LFA should be disabled.

For example, to enable TI-LFA for all IS-IS interfaces:

```
router isis 1
fast-reroute per-prefix {level-1 | level-2}
fast-reroute ti-lfa {level-1 | level-2} [maximum-metric value]
```

The **maximum-metric** option specifies the maximum repair distance which a node is still considered eligible as a release node.

To disable TI-LFA on a particular interface:

```
interface interface-name
isis fast-reroute ti-lfa protection level-1 disable
```




---

**Note** The **isis fast-reroute protection level-x** command enables local LFA and is required to enable TI-LFA.

---

2. Using interface enablement—Enable TI-LFA selectively on each interface

```

interface interface-name
isis fast-reroute protection {level-1 | level-2}
isis fast-reroute ti-lfa protection {level-1 | level-2} [maximum-metric value]

```

When both interface and protocol are TI-LFA enabled, the interface configuration takes precedence over the protocol configuration.

## Configuration Example

Example 1: In the following example, local LFA is configured with linecard-disjoint and srlg-disjoint tie-breakers. linecard-disjoint is given preference with a lower priority value (10) than the srlg-disjoint (11).

```

router isis access
net 49.0001.2037.0685.b002.00
metric-style wide
fast-flood 10
max-lsp-lifetime 65535
lsp-refresh-interval 65000
spf-interval 5 50 200
prc-interval 5 50 200
lsp-gen-interval 5 5 200
log-adjacency-changes
nsf ietf
segment-routing mpls
fast-reroute per-prefix level-1 all - configures the local LFA
fast-reroute per-prefix level-2 all
fast-reroute remote-lfa level-1 mpls-ldp - enables rLFA (optional)
fast-reroute remote-lfa level-2 mpls-ldp
fast-reroute ti-lfa level-1 - enables TI-LFA
microloop avoidance rib-update-delay 15000
bfd all-interfaces

```

Example 2—Enable TI-LFA node-protecting tie-breaker on all ISIS level-2 interfaces with priority 100. All other tie-breakers are disabled.

```

router isis
fast-reroute per-prefix level-2 all
fast-reroute ti-lfa level-2
fast-reroute tie-break level-2 node-protecting 100

```

Example 3—Enable TI-LFA node-protecting tie-breaker with priority 100 and TI-LFA SRLG protection with priority 200 on all IS-IS level-2 interfaces. All other tiebreakers are disabled because the node-protecting tie-breaker is configured.

```

router isis
fast-reroute per-prefix level-2 all
fast-reroute ti-lfa level-2
fast-reroute tie-break level-2 node-protecting 100
fast-reroute tie-break level-2 srlg-disjoint 200

```

Example 3—Enable TI-LFA node-protecting tie-breaker with priority 100 on all ISIS level-2 interfaces except on Ethernet0/0. For those IS-IS interfaces, all other tiebreakers are disabled. Ethernet0/0 overwrites the inheritance and uses the default set of tiebreakers with linecard-disjoint, lowest-backup-path-metric, srlg-disjoint enabled.

```

router isis

```

```

fast-reroute per-prefix level-2 all
fast-reroute ti-lfa level-2
fast-reroute tie-break level-2 node-protecting 100
!
interface ethernet0/0
ip router isis
isis fast-reroute tie-break level-2 default

```

**Example 4—Enable TI-LFA using the default tiebreaker on all IS-IS interfaces except on Ethernet0/0. On Ethernet0/0 enable TI-LFA node-protecting with priority 100 and disable all other tiebreakers.**

```

router isis
fast-reroute per-prefix level-2 all
fast-reroute ti-lfa level-2
!
interface ethernet0/0
ip router isis
isis fast-reroute tie-break level-2 node-protecting 100

```

**Example 5—Enable TI-LFA node-protecting tie-breaker with priority 200 and linecard-disjoint tie-breaker with priority 100 on all ISIS level-2 interfaces. All other tiebreakers are disabled.**

```

router isis
fast-reroute per-prefix level-2 all
fast-reroute ti-lfa level-2
fast-reroute tie-break level-2 linecard-disjoint 100
fast-reroute tie-break level-2 node-protecting 200

```

## Verifying the Tie-breaker

To view tiebreakers are enabled on the interface:

```

Router# show running-configuration | router isis access
Building configuration...

Current configuration : 702 bytes
!
Configuration of Partition - router isis access
!
router isis access
net 49.0001.2037.0685.b002.00
metric-style wide
fast-flood 10
max-lsp-lifetime 65535
lsp-refresh-interval 65000
spf-interval 5 50 200
prc-interval 5 50 200
lsp-gen-interval 5 5 200
no hello padding point-to-point
log-adjacency-changes
nsf cisco
nsf interval 0
segment-routing mpls
fast-reroute per-prefix level-1 all
fast-reroute per-prefix level-2 all
fast-reroute tie-break level-1 linecard-disjoint 12
fast-reroute remote-lfa level-1 mpls-ldp
fast-reroute remote-lfa level-2 mpls-ldp
fast-reroute ti-lfa level-1
bfd all-interfaces

```



```

mpls traffic-eng router-id Loopback0
mpls traffic-eng level-1
!
!
end

```

Similarly, to view the tiebreakers enabled for the router mode:

```

Router# show running-configuration | isis neighbor
Tag access:
System Id      Type Interface      IP Address      State Holdtime Circuit Id
920-CE1        L1   Gi0/2/0           1.1.1.1         UP    25      02
9k-1           L1   Gi0/2/3           14.0.0.2        UP    27      00

Router(config-srmppls)# do sh run | sec interface GigabitEthernet0/2/0
interface GigabitEthernet0/2/0
 srlg gid 5
 srlg gid 10
 ip unnumbered Loopback0
 ip router isis access
 ip ospf network point-to-point
 carrier-delay down msec 1
 negotiation auto
 ipv6 address 10:1::2/64
 mpls ip
 mpls traffic-eng tunnels
 bfd template BFD1
 cdp enable
 isis network point-to-point
903-PE1(config-srmppls)#do sh run | sec interface GigabitEthernet0/2/3
interface GigabitEthernet0/2/3
 srlg gid 10
 ip address 14.0.0.1 255.255.255.0
 ip router isis access
 ip ospf network point-to-point
 negotiation auto
 mpls ip
 mpls traffic-eng tunnels
 cdp enable
 isis circuit-type level-1
 isis network point-to-point

```

## Verifying the Primary and Repair Paths

In this example, 1.1.1.1 is the protecting neighbor and 4.4.4.4 is the neighbor on the protecting link.

```

Router#
show ip cef 1.1.1.1
1.1.1.1/32
  nexthop 1.1.1.1 GigabitEthernet0/2/0 label [explicit-null|explicit-null]() - slot 2 is
primary interface
  repair: attached-nexthop 24.0.0.2 TenGigabitEthernet0/3/0 - slot 3 is repair interface
  nexthop 24.0.0.2 TenGigabitEthernet0/3/0 label [explicit-null|explicit-null]()
  repair: attached-nexthop 1.1.1.1 GigabitEthernet0/2/0
Router#
show ip cef 4.4.4.4
4.4.4.4/32
  nexthop 4.4.4.4 GigabitEthernet0/2/3 label [explicit-null|16004]() - slot 2 is primary
interface
  repair: attached-nexthop 5.5.5.5 MPLS-SR-Tunnel2
Router# show ip cef 4.4.4.4 int

```

```

4.4.4.4/32, epoch 3, RIB[I], refcnt 6, per-destination sharing
sources: RIB, Adj, LTE
feature space:
  IPRM: 0x00028000
  Broker: linked, distributed at 4th priority
  LFD: 4.4.4.4/32 2 local labels
  dflt local label info: global/877 [0x3]
  sr local label info: global/16004 [0x1B]
    contains path extension list
    dflt disposition chain 0x46654200
      label implicit-null
      FRR Primary
        <primary: IP adj out of GigabitEthernet0/2/3, addr 4.4.4.4>
    dflt label switch chain 0x46654268
      label implicit-null
      TAG adj out of GigabitEthernet0/2/3, addr 4.4.4.4
    sr disposition chain 0x46654880
      label explicit-null
      FRR Primary
        <primary: TAG adj out of GigabitEthernet0/2/3, addr 4.4.4.4>
    sr label switch chain 0x46654880
      label explicit-null
      FRR Primary
        <primary: TAG adj out of GigabitEthernet0/2/3, addr 4.4.4.4>
subblocks:
  Adj source: IP adj out of GigabitEthernet0/2/3, addr 4.4.4.4 464C6620
  Dependent covered prefix type adjfib, cover 0.0.0.0/0
ifnums:
  GigabitEthernet0/2/3(11): 4.4.4.4
  MPLS-SR-Tunnel2(1022)
path list 3B1FC930, 15 locks, per-destination, flags 0x4D [shble, hvsh, rif, hwcn]
path 3C04D5E0, share 1/1, type attached nexthop, for IPv4, flags [has-rpr]
  MPLS short path extensions: [rib | lblmrg | srlbl] MOI flags = 0x21 label explicit-null

  nexthop 4.4.4.4 GigabitEthernet0/2/3 label [explicit-null|16004](), IP adj out of
  GigabitEthernet0/2/3, addr 4.4.4.4 464C6620
    repair: attached-nexthop 5.5.5.5 MPLS-SR-Tunnel2 (3C04D6B0)
  path 3C04D6B0, share 1/1, type attached nexthop, for IPv4, flags [rpr, rpr-only]
    MPLS short path extensions: [rib | lblmrg | srlbl] MOI flags = 0x1 label 16004
    nexthop 5.5.5.5 MPLS-SR-Tunnel2 label 16004(), repair, IP midchain out of
  MPLS-SR-Tunnel2 46CE2440
  output chain:
    label [explicit-null|16004]()
    FRR Primary (0x3B209220)
      <primary: TAG adj out of GigabitEthernet0/2/3, addr 4.4.4.4 464C6480> - primary path
      <repair: TAG midchain out of MPLS-SR-Tunnel2 46CE22A0
        label 16()
        label 16003()
        TAG adj out of TenGigabitEthernet0/3/0, addr 24.0.0.2 46CE25E0> - repair
path

```

## Verifying the IS-IS Segment Routing Configuration

```

Router# show isis segment-routing
ISIS protocol is registered with MFI
ISIS MFI Client ID:0x63
Tag Null - Segment-Routing:
  SR State:SR_ENABLED
  Number of SRGB:1
  SRGB Start:14000, Range:1001, srgb_handle:0xE0934788, srgb_state: created
  Address-family IPv4 unicast SR is configured
  Operational state: Enabled

```

The command with keyword **global-block** displays the SRGB and the range for LSPs.

```
Router# show isis segment-routing global-block
IS-IS Level-1 Segment-routing Global Blocks:
System ID          SRGB Base    SRGB Range
nevada             20000       4001
arizona            * 16000     1000
utah               40000       8000
```

The **show isis segment-routing prefix-sid-map** command with keyword **advertise** displays the prefix-sid maps that the router advertises.

```
Rrouter# show isis segment-routing prefix-sid-map adv
IS-IS Level-1 advertise prefix-sid maps:
Prefix            SID Index    Range        Flags
16.16.16.16/32    101          1
16.16.16.17/32    102          1            Attached
```

The **show isis segment-routing prefix-sid-map** command with keyword **receive** displays the prefix-sid maps that the router receives.

```
Router #sh isis segment-routing prefix-sid-map receive
IS-IS Level-1 receive prefix-sid maps:
Host              Prefix            SID Index    Range        Flags
utah              16.16.16.16/32    101          1
                  16.16.16.17/32    102          1            Attached
```

To display the connected-SIDs found in the LSPs and passed to the mapping server component, use the **show isis segment-routing connected-sid** command.

```
Router# show isis segment-routing connected-sid
IS-IS Level-1 connected-sids
Host              Prefix            SID Index    Range        Flags
nevada            * 1.1.1.2/32      1002         1
                  2.2.2.2/32        20           1
                  100.1.1.10/32     10           1
colorado          1.1.1.3/32        33           1
                  1.1.1.6/32        6            1
IS-IS Level-2 connected-sids
Host              Prefix            SID Index    Range        Flags
```

## Verifying the IS-IS TI-LFA Tunnels

```
Router# show isis fast-reroute ti-lfa tunnel
Fast-Reroute TI-LFA Tunnels:
Tunnel  Interface  Next Hop          End Point      Label          End Point Host
MP1     Et1/0         30.1.1.4          1.1.1.2        41002          nevada
MP2     Et0/0         19.1.1.6          1.1.1.6        60006          colorado
                  1.1.1.2         16             nevada
MP3     Et0/0         19.1.1.6          1.1.1.6        60006          colorado
                  1.1.1.2         16             nevada
                  1.1.1.5        70005          wyoming
```





## CHAPTER 5

# Using Segment Routing with OSPF



**Note** Using segment routing with OSPF is similar to using it with IS-IS. Before reading this chapter, see [Using Segment Routing with IS-IS](#).

Segment Routing (SR) is a new paradigm of source routing driven by the IGPs and centered on inserting a list of instructions called segments into each packet. SR forwarding can be instantiated by MPLS or IPv6.

OSPF as an IGP propagates the segment information and may either use it internally (for example, to compute Fast Rerouting repair paths) or hand over the segment information to other clients within the router (for example, TE).

To provide a base SR functionality, OSPF interacts with several components:

- Segment Routing Application—It handles IGP and instance-independent global commands as well as manages global resources used by the SR such as, block of static MPLS labels available for SR Node SID MPLS instantiation).
- MPLS Forwarding Infrastructure (MFI)—Supports MPLS forwarding table.
- RIBv4—Supports the IPv4 routing table
- [Restrictions for Using Segment Routing with OSPF, on page 39](#)
- [Enabling Segment Routing for OSPF, on page 40](#)
- [Using OSPF Fast Reroute with Segment Routing, on page 41](#)
- [Verifying the Configuration of SR on OSPF, on page 44](#)
- [Troubleshooting Segment Routing on OSPF, on page 47](#)
- [SR with OSPF Point-to-Multipoint Non Broadcast Mode, on page 47](#)
- [Additional References, on page 49](#)

## Restrictions for Using Segment Routing with OSPF

- Segment routing must be configured at the router level before enabling it at the OSPF instance.
- OSPF instance must successfully register itself with both, Segment Routing Application and MFI and retrieve resources that are necessary for operation, such as SRGB range, dynamically-assigned MPLS labels for Adjacency SIDs, and so on.
- If more than one SR router in the network advertises SID for the same prefix, then the SID values and flags must match. If conflicting attributes are found when calculating the SID for the prefix, then all SIDs for the prefix are ignored.

- Network type point-to-point and broadcast are supported.

## Enabling Segment Routing for OSPF

You can enable segment routing for IGPs under the router configuration sub mode, through commands. However, IGP segment routing are enabled only after the global SR is configured.

SR functionality configurable globally, that is outside the context of any particular IGP instance, include:

- Global enabling SR on the router.
- Specifying the range of MPLS label values to be used to instantiate SR SIDs into MPLS dataplane (SRGB block)
- Associating SID index with local prefix

## Configuring SR on an OSPF Instance

After SR is enabled globally on the router, it must be enabled in those OSPF instances that need to run SR.

**segment-routing [area *N*] {mpls | disable}**

This command is accepted only if SR is already enabled globally. If the **area** keyword is specified then SR is enabled in that area only, otherwise it is enabled in all areas attached to this OSPF instance. The keyword **disable** can be specified only when the **area** keyword is also present.

Enabling this command causes OSPF to originate RI LSA, Extended Prefix and Extended Link LSAs. It enables MPLS on all interfaces in area(s) enabled for SR and programs SR MPLS labels for forwarding.

## Enabling Advertisement of Mapping Server Prefix Ranges

Global SR configuration may contain prefix-to-SID mapping entries for prefixes that are not local to the router. Each of these entries specify the range of prefixes. Remote mapping entries can be used to find SIDs for prefixes connected to routers that do not supporting SR and hence, are not capable of advertising SIDs themselves. This is part of SR-LDP inter-working functionality.

OSPF learns the ranges configured in the global SR configuration through the SR application and advertises them in the Extended Prefix Range TLVs.

The following command allows an OSPF instance to advertise mapping entries configured in the global SR mode:

**[no] segment-routing prefix-sid-map advertise-local**

This command is configurable in the router mode. Its default state is *disabled* , That is, no mapping ranges are advertised by OSPF even if they are configured in the global SR mode.

## Disabling Mapping Server Functionality

By default, OSPF considers prefix ranges received from mapping servers in the network when computing SIDs for prefixes. This is the core of SR-LDP inter-working label computation. However, if the SR-LDP feature needs to be disabled, use the following command:

**[no] segment-routing prefix-sid-map receive**

Default state of this command is *enabled*. That is, OSPF processes mapping ranges received from mapping servers in the network.

When this command is configured in its non-default form, or the **no** form, it prohibits OSPF from considering prefix-SID mappings from mapping servers.



**Note** This command does not affect processing of 'native' SIDs, that is, those SIDs that are advertised in the Extended Prefix TLVs by routers to whom the prefix is locally connected.

## Interface Mode Prefix Attributes Command

Extended Prefix TLV of the Extended Prefix LSA carries flags for the prefix and one of them is N-flag (Node). The N-flag indicates that any traffic sent along to the prefix is destined to the router originating the LSA. This flag typically marks the host routes of the router's loop-back.

By default, OSPF advertises the routes of its loopback with the N-flag. If the IP address of a loopback is not unique (for example, it is part of an anycast address), then OSPF must be configured to *not* advertise the N-flag with the prefix.

Use the following command to cause OSPF to clear N-flag on all prefixes configured on interface in the interface-mode:

**ip ospf prefix-attributes n-flag-clear**

## Using OSFP Fast Reroute with Segment Routing

IP Fast Reroute is a set of techniques that allow rerouting IP traffic around a failed link or a failed node in the network within a very short time (< 50ms). One of the techniques to do this is Loop Free Alternates (LFA). Effective with Cisco IOS XE Release 3.18S, OSPF supports per-prefix directly connected LFA and remote LFA (RLFA).

The per-prefix directly connected LFA provides loop-free alternate path for most triangular topologies, but does not provide good coverage for rectangular or circular topologies. However, the RLFA, which uses MPLS forwarding with LDP signaling for tunneling the rerouted traffic to an intermediate node, extends the IPFRR coverage in ring or rectangular topologies. For each link, RLFA defines the P-Space (the set of nodes reachable from the calculating node without crossing the protected link) and Q-Space (the set of nodes that can reach the neighbor on the protected link without crossing the protected link itself). The nodes that belong to both P- and Q-Spaces are called PQ nodes and can be used as the intermediate nodes for the protected traffic. However, for topologies where the P- and Q-Spaces are completely disjoint, there is still no coverage by RLFA.

Topology Independent Fast Reroute (TI-FRR) is a technique that uses Segment Routing to provide link protection in any topology, assuming the metric on the links in the topology is symmetrical. Even TI-LFA does not guarantee a backup in cases where the bandwidth on a single link is asymmetrical. TI-LFA only considers loop-free paths that are on the post-convergence path that helps you to better plan the capacity of the network.

Segment Routing allows creating a full explicit path through the network, but using such a fully specified path is not scalable in larger topologies due to the number of segments along the path. Specifying the whole path is, however, not necessary, and only a subset of the path is needed to carry the traffic to an intermediate

node (release node) which does not loop the traffic back to the protecting node. When the release node is one of the neighbors of S (source or sender node), then we have a directly connected TI-LFA.

Effective with Cisco IOS XE Release 3.18S, an SR Tunnel is constructed to the release node (if remote) and used as the TI-LFA. The tunnel is constructed by explicitly forwarding through a set of one or more repair nodes. The tunnel is created using a SID stack consisting of one SID (Node SID or Adjacency SID) for every repair node and the corresponding label stack is pushed to the protected traffic.

## Per Instance Enablement

- TI-LFA can be enabled on a per interface basis. One interface could be configured for LFA, another interface for LFA and RLFA, yet another interface for LFA and TI-LFA and another interface for LFA, RLFA, and TI-LFA.
- TI-LFA backup path is calculated only if TI-LFA protection is enabled on the output interface of the primary path.
- TI-LFA protection is not available for virtual links, sham links, and TE tunnels.
- Backup paths do not use virtual links, sham links, and TE tunnels.

## Limitations of TI-LFA on OSPF

- TI-LFA is supported only on OSPFv2.
- TI-LFA only computes TI-LFA candidates that have been found based on a post-convergence path excluding the primary link. That is, TI-LFA only provides link protection. Node protection or SRLG protection is not supported.
- TI-LFA is calculated only if D and repair nodes (P, Q, PQ, or other) in the post-convergence path are segment routing capable. S does not need to be SR capable.
- TI-LFA is calculated only if Repair Nodes (P, Q, PQ, or other) in the post-convergence path have node SID. A node SID can be directly connected or one advertised by Mapping Server (SRMS).
- TI-LFA is restricted to a maximum of two or three segment tunnels, which may be lesser than the maximum labels that can be pushed by the router.
- When S is not SR capable, TI-LFA calculates a 0 or 1 segment LFA along the post-convergence path. The 1-segment LFA uses the LDP label to reach the PQ node in the post-convergence path.
- TI-LFA does not compute backup for virtual link, sham link, or TE tunnel path.
- TI-LFA does not use virtual link, sham link or TE tunnel as backup path.
- TI-LFA does not provide adjacency protection. It only protects prefixes. SR TE may provide protection for adjacency segments. Packets arriving with active adjacency segment do not have TI-LFA protection.
- TI-LFA is calculated only for prefixes having SR prefix SID or node SID.
- TI-LFA does not support Multi Topology Routing (MTR). LFA is calculated only for MTID zero.
- Cisco ASR900 routers with RSP3 module, support a maximum of two labels under TI-LFA tunnel.



## Enabling and Disabling TI-FRR

TI-LFA can be enabled on interface, area, and process scopes. By default, TI-LFA is disabled globally.

TI-LFA can be enabled using following commands.

### **fast-reroute per-prefix ti-lfa**

However, when TI-LFA is enabled globally, you have the option to disable it per area level, using following command:

**[no] fast-reroute per-prefix ti-lfa [area *area* disable]**

This command is accepted even if Segment-Routing is not enabled on the OSPF instance or area. In this case, a TI-LFA backup path is created if the post-convergence path has a PQ node with an LDP label to reach it. That is, RLFA through a post-convergence path.

At the process level:

- **fast-reroute per-prefix ti-lfa [area *area* disable]**
- **fast-reroute per-prefix enable [area *area*] prefix-priority {high | low}**
- **fast-reroute per-prefix remote-lfa [area *area*] maximum-cost *cost***
- **fast-reroute per-prefix remote-lfa [area *area*] tunnel mpls-ldp**
- **fast-reroute per-prefix remote-lfa [area *area*] tunnel mpls-ldp**
- **fast-reroute per-prefix tie-break *attribute* [required] index *index***

At the interface level:

- **ip ospf fast-reroute per-prefix protection disabled**
- **ip ospf fast-reroute per-prefix candidate disable**
- **ip ospf fast-reroute per-prefix protection ti-lfa [disable]**

## Verifying TI-LFA on OSPF

The following show command displays the details about TI-LFA configuration:

### **show ip ospf fast-reroute ti-lfa**

```
OSPF Router with ID (13.13.13.13) (Process ID 100)
Microloop avoidance is enabled for protected prefixes, delay 5000 msec
Loop-free Fast Reroute protected prefixes:
Area Topology name Priority Remote LFA Enabled TI-LFA Enabled
0 Base High Yes Yes
Repair path selection policy tiebreaks (built-in default policy):
0 post-convergence
10 primary-path
20 interface-disjoint
30 lowest-metric
40 linecard-disjoint
50 broadcast-interface-disjoint
256 load-sharing
OSPF/RIB notifications:
Topology Base: Notification Enabled, Callback Registered
Last SPF calculation started 00:02:42 ago and was running for 8 ms.
```

## Verifying the Configuration of SR on OSPF

- The following **show** command displays information about the new LSAs, such as router information, Extended Prefix and Extended Link LSAs:

### **show ip ospf database opaque-area type ext-link**

```

OSPF Router with ID (13.13.13.13) (Process ID 100)
Type-10 Opaque Area Link States (Area 0)
LS age: 1378
Options: (No TOS-capability, DC)
LS Type: Opaque Area Link
Link State ID: 8.0.0.7
Opaque Type: 8 (Extended Link)
Opaque ID: 7
Advertising Router: 4.4.4.4
LS Seq Number: 80000056
Checksum: 0xFE8
Length: 76
TLV Type: Extended Link
Length: 52
Link connected to : another Router (point-to-point)
(Link ID) Neighboring Router ID: 12.12.12.12
(Link Data) Interface IP address: 0.0.0.7
Sub-TLV Type: Adj SID
Length : 7
Flags : L-Bit, V-bit
MTID : 0
Weight : 0
Label : 27
Sub-TLV Type: Adj SID
Length : 7
Flags : L-Bit, V-bit, B-bit
MTID : 0
Weight : 0
Label : 34
Sub-TLV Type: Local / Remote Intf ID
Local Interface ID : 7
Remote Interface ID : 0

```

### **show ip ospf database opaque-area type ext-prefix**

```

OSPF Router with ID (13.13.13.13) (Process ID 100)
Type-10 Opaque Area Link States (Area 0)
LS age: 1392
Options: (No TOS-capability, DC)
LS Type: Opaque Area Link
Link State ID: 7.0.0.0
Opaque Type: 7 (Extended Prefix)
Opaque ID: 0
Advertising Router: 4.4.4.4
LS Seq Number: 80000055
Checksum: 0xD156
Length: 44
TLV Type: Extended Prefix
Length: 20
Prefix : 4.4.4.4/32
AF : 0
Route-type: Intra
Flags : N-bit

```

```

Sub-TLV Type: Prefix SID
Length: 8
Flags : None
MTID : 0
Algo : IGP metric based SPT
SID : 4

```

### **show ip ospf database opaque-area type router-information**

```

OSPF Router with ID (13.13.13.13) (Process ID 100)
Type-10 Opaque Area Link States (Area 0)
LS age: 1402
Options: (No TOS-capability, DC)
LS Type: Opaque Area Link
Link State ID: 4.0.0.0
Opaque Type: 4 (Router Information)
Opaque ID: 0
Advertising Router: 4.4.4.4
LS Seq Number: 80000055
Checksum: 0x7B86
Length: 52
TLV Type: Router Information
Length: 4
Capabilities:
Graceful Restart Helper
Stub Router Support
Traffic Engineering Support
TLV Type: Segment Routing Algorithm
Length: 1
Algorithm: IGP metric based SPT
TLV Type: Segment Routing Range
Length: 12
Range Size: 8000
Sub-TLV Type: SID/Label
Length: 3
Label: 16000

```

- The following show command displays the interface output, but only if segment routing auto-enabled MPLS forwarding on the interface and SIDs were allocated to that interface

### **show ip ospf interface**

```

Loopback0 is up, line protocol is up
Internet Address 13.13.13.13/32, Interface ID 29, Area 0
Attached via Interface Enable
Process ID 100, Router ID 13.13.13.13, Network Type LOOPBACK, Cost: 1
Topology-MTID Cost Disabled Shutdown Topology Name
0 1 no no Base
Enabled by interface config, including secondary ip addresses
Loopback interface is treated as a stub Host
BDI2104 is up, line protocol is up
Interface is unnumbered, Interface ID 40, Area 0
Using address of Loopback0 (13.13.13.13)
Attached via Interface Enable
Process ID 100, Router ID 13.13.13.13, Network Type POINT_TO_POINT, Cost: 1
Topology-MTID Cost Disabled Shutdown Topology Name
0 1 no no Base
Enabled by interface config, including secondary ip addresses
Transmit Delay is 1 sec, State POINT_TO_POINT, BFD enabled
Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
oob-resync timeout 40
Hello due in 00:00:05
Supports Link-local Signaling (LLS)

```

```

Cisco NSF helper support enabled
IETF NSF helper support enabled
Can be protected by per-prefix Loop-Free FastReroute
Can be used for per-prefix Loop-Free FastReroute repair paths
Not Protected by per-prefix TI-LFA
Segment Routing enabled for MPLS forwarding
Index 1/3/3, flood queue length 0
Next 0x0(0)/0x0(0)/0x0(0)
Last flood scan length is 1, maximum is 2
Last flood scan time is 0 msec, maximum is 0 msec
Neighbor Count is 1, Adjacent neighbor count is 1
Adjacent with neighbor 8.8.8.8
Suppress hello for 0 neighbor(s)

```

- The following show command displays the information about segment routing
  - Per-instance state of the feature—whether enabled or disabled, whether operational or not
  - Registration status of OSPF with the SR application and MFI
  - Retrieved SRGB range

### **show ip ospf [AS] segment-routing**

```

OSPF Router with ID (13.13.13.13) (Process ID 100)
Global segment-routing state: Enabled
Segment Routing enabled:
Area Topology name Forwarding
0 Base MPLS
SR Attributes
Prefer SR Labels
Do not advertise Explicit Null
Local MPLS label block (SRGB):
Range: 16000 - 23999
State: Created
Registered with SR App, client handle: 10
Connected map notifications active (handle 0x18), bitmask 0x1
Active policy map notifications active (handle 0x19), bitmask 0x0
Registered with MPLS, client-id: 100
Bind Retry timer not running
Adj Label Bind Retry timer not running
Adj Protected Label Bind Retry timer not running

```

- The following show command displays information about the SID database for one or for all SIDs. The output displays all prefixes that were advertised with a given SID and highlights which prefixes are local to the router.



#### **Note**

In the absence of a mis-configuration, there is only one prefix reported for each SID value.

### **show ip ospf [AS] segment-routing sid-database [SID]**

```

OSPF Router with ID (13.13.13.13) (Process ID 100)
OSPF Segment Routing SIDs
Flags: L - local, N - label not programmed,
M - mapping-server
SID Prefix/Mask Adv-Rtr-Id Area-Id
-----
8 8.8.8.8/32 8.8.8.8 0
10 10.10.10.10/32 10.10.10.10 0
13 (L) 13.13.13.13/32 13.13.13.13 0

```

- The following show command displays the database of local prefixes along with their SID values and flags.

**Note**

This show command provides information only about locally-configured prefixes. Information about prefixes received from other routers is available either through LSDB (**show ip ospf database opaque-area**) or LRIB (**show ip ospf rib local**) show commands.

**show ip ospf [AS] segment-routing local-prefix**

```
OSPF Router with ID (13.13.13.13) (Process ID 100)
Area 0:
Prefix: Sid: Index: Interface:
13.13.13.13/32 13 0.0.0.0 Loopback0
```

## Troubleshooting Segment Routing on OSPF

Use the following command to debug issues

- **debug ip ospf [AS] segment-routing**
- **debug ip ospf fast-reroute spf**
- **debug ip ospf fast-reroute spf detail**
- **debug ip ospf fast-reroute rib**
- **debug ip ospf fast-reroute rib [access-list]**

## SR with OSPF Point-to-Multipoint Non Broadcast Mode

*Table 1: Feature History Table*

Feature Name	Release Information	Description
SR with OSPF Point-to-Multipoint Non Broadcast Mode	Cisco IOS XE Cupertino 17.7.1	This feature enables SR on Point-to-Multipoint (P2MP) Non Broadcast (NB) OSPF network type, where you can manually configure neighbor IP address and OSPF sends hello packet using the neighbor unicast address.

The following are the three major network types defined in OSPF:

### 1. Point-to-Point Network

This is the simplest form of the network types. Two routers are connected together with a single link. An example of a point-to-point link is a serial link connecting just two routers (using HDLC or PPP). With point-to-point links, OSPF does not select a Designated Router (DR) or Backup Designated Router (BDR). This network type supports auto neighbor discovery.

### 2. Broadcast Network

In this network type, a broadcast or multicast packet sent by one router on the network can be received by all or some of the other routers. This type of network is also multi-access because more than two routers can be connected to them. Ethernet is an example of a broadcast network. DR and BDR are elected on this network.

### 3. Point-to-Multipoint (P2MP) Non Broadcast (NB) Network

This network type supports multiple routers (multi-access) but does not support the broadcast capability. In this mode, OSPF treats the non-broadcast network like a collection of point-to-point links. There is no DR/BDR election, but neighbors may be automatically discovered, depending on how the interface is configured.

Prior to Cisco IOS XE Cupertino Release 17.7.1, you could only configure SR for point-to-point broadcast network. Starting with Cisco IOS XE Cupertino Release 17.7.1, you can configure SR with P2MP NB mode, where you can manually configure neighbor IP address. In this configuration, OSPF sends hello packet using the neighbor unicast address. Another advantage of running OSPF in this mode is that a common subnet can still be used among the routers, unlike different subnets in the case of point-to-point links.

You must use SDM template, `sr_5_label_push_enable`, for increased number of SR labels. This is applicable to ASR RSP3 module.

## License Requirement

You must use metro aggregation services license to enable this feature:

```
PE1(config)#license boot level metroaggrservices
```

## Limitation

- Dynamic SRTE, SR micro loop avoidance, and Topology-Independent Loop-Free Alternate (TI-LFA) are not supported on node with P2MP NB OSPF and SR.
- Only native SR is supported.

## Configure SR with OSPF P2MP NB Network

### 1. Configure SR Block

```
segment-routing mpls
  set-attributes
  address-family ipv4
    sr-label-preferred
    explicit-null
  exit-address-family
connected-prefix-sid-map
  address-family ipv4
    Prefix SID/32 index index range 1
  exit-address-family
```

### 2. Configure OSPF

```
router ospf ospf number
  router-id RID
  segment-routing mpls
  segment-routing prefix-sid-map advertise-local
  distribute link-state instance-id unique number
```

```
neighbor neighbor interface IP
neighbor neighbor interface IP
```

### 3. Configure on Interface

```
interface interface
ip ospf network point-to-multipoint non-broadcast
ip ospf ospf number area 0
```

## Verification of SR with OSPF P2MP NB Network Configuration

Use `show ip ospf interface br` and `show ip ospf neighbor` commands to verify SR with OSPF P2MP NB mode configuration.

```
PE12#show ip ospf interface br
```

Interface	PID	Area	IP Address/Mask	Cost	State	Nbrs	F/C
Lo1	10	0	12.12.12.12/32	1	LOOP	0/0	
Fo0/5/1	10	0	192.148.1.2/24	1	P2MP	1/1	

```
PE12#show ip ospf neighbor
```

Neighbor ID	Pri	State	Dead Time	Address	Interface
13.13.13.13	0	FULL/ -	00:01:36	192.148.1.1	FortyGigabitEthernet0/5/1

## Additional References

### Related Documents

Related Topic	Document Title
Cisco IOS commands	<a href="https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/mcl/allreleasemcl/all-book.html">https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/mcl/allreleasemcl/all-book.html</a>

### Standards and RFCs

Standard/RFC	Title
No specific Standards and RFCs are supported by the features in this document.	

### MIBs

MB	MIBs Link
—	To locate and download MIBs for selected platforms, Cisco IOS releases, and feature sets, use Cisco MIB Locator found at the following URL: <a href="http://www.cisco.com/go/mibs">http://www.cisco.com/go/mibs</a>

**Technical Assistance**

Description	Link
<p>The Cisco Support website provides extensive online resources, including documentation and tools for troubleshooting and resolving technical issues with Cisco products and technologies.</p> <p>To receive security and technical information about your products, you can subscribe to various services, such as the Product Alert Tool (accessed from Field Notices), the Cisco Technical Services Newsletter, and Really Simple Syndication (RSS) Feeds.</p> <p>Access to most tools on the Cisco Support website requires a Cisco.com user ID and password.</p>	<p><a href="http://www.cisco.com/cisco/web/support/index.html">http://www.cisco.com/cisco/web/support/index.html</a></p>





## CHAPTER 6

# Segment Routing Traffic Engineering

Segment Routing (SR) enables any network node, such as a Server, Provider Edge (PE), Aggregator, or Provider (P) to engineer an explicit path for each of its traffic classes.



### Note

This explicit path does not depend on a hop-by-hop signaling technique, such as Label Distribution Protocol (LDP) or Resource Reservation Protocol (RSVP); it only depends on a set of "segments" that are preprogrammed and advertised by the link-state routing protocol.

These segments act as topological sub-paths that can be combined together to form the desired path. In Segment Routing, the path is encoded in each packet itself in the form of SR Segment Identifiers (SIDs).

There are two types of segments—prefix and adjacency.

- A prefix segment represents the shortest path (as computed by IGP) to reach a specific prefix; a node segment is a special prefix segment that is bound to the loopback address of a node.
- An adjacency segment represents a specific adjacency to a neighbor node.

A node segment can be a multi-hop path while an adjacency segment is a one-hop path.

A segment is represented by a 32-bit entity called Segment ID (SID). A prefix-SID is globally unique, and the operator ensures such uniqueness. An adjacency SID is locally unique to the node, and is automatically generated by the node attached to the adjacency.

The Segment Routing control-plane can be applied to the MPLS data-plane. In this case, the prefix-SID in the MPLS data-plane is represented as an LSP whose path flows along the shortest-path to the prefix, whereas an adjacency-SID is represented as cross-connect entry pointing to a specific egress data-link.

A Traffic Engineered (TE) tunnel is a container of TE LSPs instantiated between the tunnel ingress and the tunnel destination. A TE tunnel can instantiate one or more SR-TE LSPs that are associated with the same tunnel. The SR-TE LSP path may not necessarily follow the same IGP path to a destination node. In this case, the SR-TE path can be specified through either a set of prefix-SIDs, or adjacency-SIDs of nodes, or both, and links to be traversed by the SR-TE LSP.

The headend imposes the corresponding MPLS label stack on outgoing packets to be carried over the tunnel. Each transit node along the SR-TE LSP path uses the incoming top label to select the next-hop, pop or swap the label, and forward the packet to the next node with the remainder of the label stack, until the packet reaches the ultimate destination.

The set of hops or segments that define an SR-TE LSP path are provisioned by the operator.

- [Restrictions for SR-TE, on page 52](#)
- [Segment Routing Recirculation for TI-LFA, on page 53](#)
- [Statistics Counters on Cisco RSP3 , on page 55](#)
- [Configuring SR-TE Tunnels, on page 55](#)
- [Verifying SR-TE Tunnels, on page 56](#)
- [Verifying Statistics Counters, on page 56](#)
- [Configuring a Path Option for a TE Tunnel, on page 57](#)
- [Configuring SR Explicit Path Hops, on page 57](#)
- [Use Case: Segment Routing Traffic Engineering Basic Configuration, on page 58](#)
- [Verifying Configuration of the SR-TE Tunnels, on page 60](#)
- [SR-TE LSP Instantiation, on page 63](#)
- [SR-TE and TI-LFA, on page 71](#)

## Restrictions for SR-TE

- In the default SDM template, Cisco ASR 900 Router with RSP3 module can push a maximum of 4 MPLS labels in the egress direction. This includes service labels (L3VPN, L2VPN, 6PE/6VPE), RFC 3107 BGP-LU label and SR labels for FRR primary or backup paths. In default template, packet recirculation within forwarding ASIC is not used.
- Effective Cisco IOS XE Gibraltar 16.11.1, Cisco ASR 900 Router with RSP3 module can push a maximum of seven MPLS labels in the egress direction when segment routing traffic engineering is being used. This is supported through the template **sr\_5\_label\_push\_enable**. This template uses packet recirculation within forwarding ASIC. For more information, see [Support for Seven SR-TE Lables, on page 54](#).
- The routers do not support unequal load balancing when using the load-share option.
- ECMP at single SR-TE tunnel level is not supported.
- SR-TE FRR with PoCH as the primary path is not supported. However, you can provision SR-TE tunnel without FRR over PoCH and backup tunnel over PoCH. Ensure that the **min-link** value configured is equal to the actual number of member links.
- The SR-TE dynamic tunnels do not support node protection. Therefore, node protection cannot achieve less than 50ms convergence with dynamic SR-TE tunnels.
- Limitations for SR-TE statistic counters in Cisco IOS XE Bengaluru 17.5.1:
  - For Cisco ASR 900 Router with RSP3 module, counters on traffic engineering policies or tunnels to the next hops are not supported by default. To enable the counters for single-hop tunnels, you must configure the following CLI command:  
**platform mpls traffic-engg single-hop-tunnel-statistics-enable**  
For more information, see [Statistics Counters on Cisco RSP3 , on page 55](#).
  - Statistic counters for ECMP to first hop node is not supported.
  - Statistic counters for traffic steered over PFP policy is not supported.
  - Starting from Cisco IOS XE Release 17.5.1, statistic counters for Labelled Traffic over PDP SR policy is supported. See the [Feature History, on page 1](#) for more information.

- For L2VPN prefixes going over more than one auto-route tunnel, an ECMP LB path is formed with SR-TE. Since this is a preselection, only one path is picked up for L2VPN.
- Per-traffic class aggregate counters per-SR policy are not supported.
- Per-binding SID aggregate counters per SR-policy are not supported.
- Multiple segment list is not supported.
- Multiple segment lists and tunnels pointing to ECMP next-hop are not supported.

## Segment Routing Recirculation for TI-LFA

This increase in the number of labels is achieved by recirculating the FRR backup path.

The following table indicates the supported label combinations:

Service	Transport	TI-LFA	Recirculation in Backup
0	3	0	NA
0	2	1	Yes
0	2	2	Yes
0	1	1	No
0	1	2	Yes
1	3	0	NA
1	2	1	Yes
1	2	2	Yes
1	1	1	No
1	1	2	Yes
2	2	1	Yes
2	2	2	Yes
2	1	1	No
2	1	2	Yes

## Restrictions

- For all unsupported OCE chains, a download error message is displayed.
- PIC core is not supported over SR-TE tunnel PIC edge is not supported.

- Effective Cisco IOS XE Bengaluru 17.5.1, ECMP over SR-TE tunnels is supported on RSP3. See [ECMP Over SR-TE](#) for more details.
- Three transport labels (without TI-LFA) are not supported
- Segment routing in Active/Active QoS template is not supported

## Support for Seven SR-TE Lables

Effective Cisco IOS XE Gibraltar 16.11.1, the Cisco ASR900 routers with RSP3 module support up to four transport labels along with one to two service labels. The following table indicates the number and type of labels that can be used:

Service	Transport	TI-LFA	Recirculation in Backup
0	4	1	Yes
0	4	2	Yes
0	3	1	No
0	3	2	Yes
0	2	1	Yes
0	2	2	Yes
0	1	1	No
0	1	2	Yes
1	4	1	Yes
1	4	2	Yes
1	3	1	No
1	3	2	Yes
1	2	1	Yes
1	2	2	Yes
1	1	1	No
1	1	2	Yes
2	4	NA	Not Supported
2	3	NA	Not Supported
2	2	1	No
2	2	2	Yes
2	1	1	No

Service	Transport	TI-LFA	Recirculation in Backup
2	1	2	Yes
3	1	NA	Not supported

## Statistics Counters on Cisco RSP3

For Cisco ASR 900 Router with RSP3 module, counters on traffic engineering policies or tunnels to the next hops are not supported by default. To enable the counters for single-hop tunnels, you must configure the following CLI command:

**platform mpls traffic-engg single-hop-tunnel-statistics-enable**

You can disable the counters support for single-hop tunnels by using the "no" form of the command as follows:

**no platform mpls traffic-engg single-hop-tunnel-statistics-enable**

Any configuration change will take effect only on the new tunnel that becomes operational after the CLI command is executed. If you want the single-hop tunnel statistics to be supported on the existing tunnels, you should unconfigure and reconfigure the tunnel. For ODN tunnels, the prefixes have to be withdrawn and readvertised.

The CLI command does not support enabling or disabling the counters per tunnel.

## Configuring SR-TE Tunnels

### Configuration of Exp path

```
enable
Router1(cfg-ip-expl-path)# index 1 next-label 16003
Explicit Path name Router1-Router3-AG1:
  1: next-label 16003
Router1(cfg-ip-expl-path)# index 2 next-label 16006
Explicit Path name Router1-Router3-AG1:
  1: next-label 16003
  2: next-label 16006
Router1(cfg-ip-expl-path)# end
```

### Tunnel configuration:

```
config terminal
Enter configuration commands, one per line. End with CNTL/Z.
Router1(config)# interface Tunnell
Router1(config-if)# ip unnumbered Loopback1
Router1(config-if)# tunnel destination 6.6.6.6
Router1(config-if)# tunnel mode mpls traffic-eng
Router1(config-if)# tunnel mpls traffic-eng autoroute announce
Router1(config-if)# tunnel mpls traffic-eng path-option 1 explicit name Router1-Router3-AG1
segment-routing
Router1(config-if)# end
Router#
```

## Verifying SR-TE Tunnels

```
Router1# show mpls traffic-eng tunnel tun1

Name: CS1_t1                                     (Tunnel1) Destination: 6.6.6.6
Status:
  Admin: up          Oper: up          Path: valid          Signalling: connected
  path option 1, (SEGMENT-ROUTING) type explicit CS1-CS3-AG1 (Basis for Setup)

Config Parameters:
  Bandwidth: 0          kbps (Global)  Priority: 7 7      Affinity: 0x0/0xFFFF
  Metric Type: TE (default)
  Path Selection:
    Protection: any (default)
  Path-selection Tiebreaker:
    Global: not set    Tunnel Specific: not set    Effective: min-fill (default)
  Hop Limit: disabled [ignore: Explicit Path Option with all Strict Hops]
  Cost Limit: disabled
  Path-invalidation timeout: 10000 msec (default), Action: Tear
  AutoRoute: enabled  LockDown: disabled Loadshare: 0 [0] bw-based
  auto-bw: disabled
  Fault-OAM: disabled, Wrap-Protection: disabled, Wrap-Capable: No
Active Path Option Parameters:
  State: explicit path option 1 is active
  BandwidthOverride: disabled LockDown: disabled Verbatim: disabled

History:
  Tunnel:
    Time since created: 1 minutes, 2 seconds
    Time since path change: 47 seconds
    Number of LSP IDs (Tun_Instances) used: 14
    Current LSP: [ID: 14]
    Uptime: 47 seconds
    Prior LSP: [ID: 1]
    ID: path option unknown
    Removal Trigger: unspecified
  Tun_Instance: 14
  Segment-Routing Path Info (isis level-1)
    Segment0[Node]: 3.3.3.3, Label: 16003
    Segment1[Node]: 6.6.6.6, Label: 16006
Router#
```

## Verifying Statistics Counters

The following sample output shows how the statistics counters are displayed:

```
router#show segment-routing traffic-eng policy name R1_R5_TE detail

Name: R1_R5_TE (Color: 5 End-point: 5.5.5.5)
Owners : CLI
Status:
  Admin: up, Operational: down for 337:42:13 (since 01-04 16:36:35.093)
Candidate-paths:
  Preference 150 (CLI):
    Dynamic (pce) (inactive)
      Inactive Reason: PCE reply is pending
      Metric Type: TE
Attributes:
  Binding SID: 22
```

```

Allocation mode: dynamic
State: Programmed
Tunnel ID: 65541 (Interface Handle: 0x41)
Per owner configs:
  CLI
  Binding SID: dynamic
Stats:
  5 minute output rate 0 bits/sec, 0 packets/sec
  Packets: 1800      Bytes: 1092      <<< packets that are traversing via tunnel are
shown>>>
PM profile: Not configured

```

## Configuring a Path Option for a TE Tunnel

The **segment-routing** keyword indicates that the specified path is programmed as an SR path.

```

Router(config)# interface tunnel 100
Router(config-if)# tunnel mpls traffic-eng path-option 1 explicit name foo segment-routing
Router(config-if)# tunnel mpls traffic-eng path-option 2 dynamic segment-routing
Router(config-if)# tunnel mpls traffic-eng path-option 3 segment-routing

```

When the path-option type for an operational SR tunnel is changed from SR to non-SR (for example, **dynamic**), the existing forwarding entry of the tunnel is deleted.

Segment Routing can be enabled or disabled on an existing secondary or an in-use path-option. If the tunnel uses a signaled RSVP-TE explicit path-option and segment routing is enabled on that tunnel, the RSVP-TE LSP is torn, and the SR-TE LSP is instantiated using the same path-option. Conversely, if segment routing is disabled on a path-option that is in use by the primary LSP, the tunnel goes down intermittently and a new RSVP-TE LSP is signaled using the same explicit path.

If the “segment-routing” path-option is enabled on a secondary path-option (that is, not in use by the tunnel’s primary LSP), the tunnel is checked to evaluate if the newly specified SR-TE LSP path-option is valid and more favorable to use for the tunnel primary LSP.

## Configuring SR Explicit Path Hops

For intra-area LSPs, the explicit path can be specified as a list of IP addresses:

```

Router(config)# ip explicit-path name foo
Router(config-ip-expl-path)# index 10 next-address 1.1.1.1 --> node address
Router(config-ip-expl-path)# index 20 next-address 12.12.12.2 --> link address

```

The explicit path can also be specified as segment-routing SIDs:

```

(config)# ip explicit-path name foo
(config-ip-expl-path)# index 10 next-label 20

```

The following SR-TE explicit path hops are supported:

- IP addresses
- MPLS labels

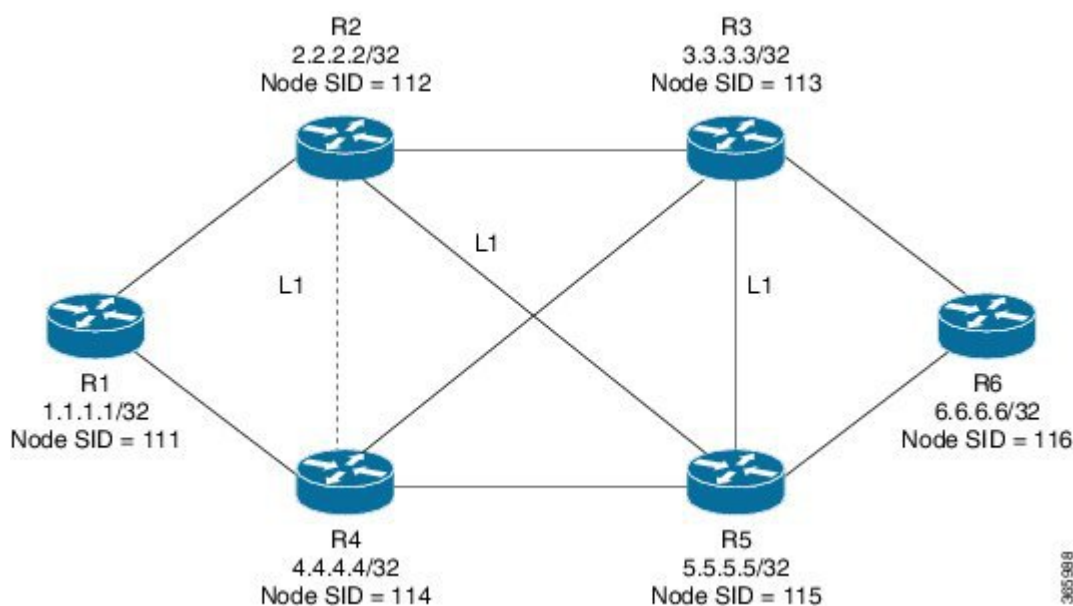
- Mix of IP addresses and MPLS labels



**Note** IP addresses cannot be used after using the label in MIXED\_PATH.

## Use Case: Segment Routing Traffic Engineering Basic Configuration

Consider the following topology:



Configuration at the headend router, R1:

```
interface GigabitEthernet0/02
ip address 100.101.1.1 255.255.255.0
ip router isis 1
isis network point-to-point
negotiation auto
mpls traffic-eng tunnels
router isis 1
net 49.0001.0010.0100.1001.00
is-type level-1
ispf level-1
metric-style wide
log-adjacency-changes
segment-routing mpls
segment-routing prefix-sid-map advertise-local
mpls traffic-eng router-id Loopback1
mpls traffic-eng level-1
```

Configuration at the tailend router, R6



```
interface GigabitEthernet0/02
ip address 100.101.1.1 255.255.255.0
ip router isis 1
isis network point-to-point
negotiation auto
mpls traffic-eng tunnels
router isis 1
 net 49.0001.0060.0600.6006.00
 ispf level-1
 metric-style wide
 log-adjacency-changes
 segment-routing mpls
segment-routing prefix-sid-map advertise-local
 mpls traffic-eng router-id Loopback1
 mpls traffic-eng level-1
```

## Explicit Path SR-TE Tunnel 1

Consider tunnel 1 based only on IP addresses:

```
ip explicit-path name IP_PATH1
 next-address 2.2.2.2
 next-address 3.3.3.3
 next-address 6.6.6.6
!
interface Tunnel1
 ip unnumbered Loopback1 poll point-to-point
 tunnel mode mpls traffic-eng
 tunnel destination 6.6.6.6
 tunnel mpls traffic-eng autoroute announce
 tunnel mpls traffic-eng priority 6 6
 tunnel mpls traffic-eng path-option 10 explicit name IP_PATH1 segment-routing
 tunnel mpls traffic-eng path-selection metric igp
 tunnel mpls traffic-eng 10
end
```

## Explicit Path SR-TE Tunnel 2

Consider tunnel 2 based on node SIDs

```
ip explicit-path name IA_PATH
 next-label 114
 next-label 115
 next-label 116
!
interface Tunnel2
 ip unnumbered Loopback1 poll point-to-point
 tunnel mode mpls traffic-eng
 tunnel destination 6.6.6.6
 tunnel mpls traffic-eng autoroute announce
 tunnel mpls traffic-eng priority 6 6
 tunnel mpls traffic-eng bandwidth 10000 class-type 1
 tunnel mpls traffic-eng path-option 10 explicit name NODE_PATH segment-routing
 tunnel mpls traffic-eng path-selection metric igp
 tunnel mpls traffic-eng 10
end
```

## Explicit Path SR-TE Tunnel 3

Consider that tunnel 3 is based on a mix of IP addresses and label

```
ip explicit-path name MIXED_PATH enable
next-address 2.2.2.2
next-address 3.3.3.3
next-label 115
next-label 116
!
interface Tunnel3
ip unnumbered Loopback1 poll point-to-point
tunnel mode mpls traffic-eng
tunnel destination 6.6.6.6
tunnel mpls traffic-eng autoroute announce
tunnel mpls traffic-eng priority 6 6
tunnel mpls traffic-eng path-option 10 explicit name MIXED_PATH segment-routing
tunnel mpls traffic-eng path-selection metric igp
tunnel mpls traffic-eng 10
```

## Dynamic Path SR-TE Tunnel 4

Consider that tunnel 4 is based on adjacency SIDs

```
interface Tunnel4
ip unnumbered Loopback1 poll point-to-point
tunnel mode mpls traffic-eng
tunnel destination 6.6.6.6
tunnel mpls traffic-eng autoroute announce
tunnel mpls traffic-eng priority 6 6
tunnel mpls traffic-eng bandwidth 10000 class-type 1
tunnel mpls traffic-eng path-option 10 dynamic segment-routing
tunnel mpls traffic-eng path-selection metric igp
tunnel mpls traffic-eng 10
end
```

## Dynamic Path SR-TE Tunnel 5

Consider that tunnel 5 is based on Node SIDs

```
interface Tunnel5
ip unnumbered Loopback1 poll point-to-point
tunnel mode mpls traffic-eng
tunnel destination 6.6.6.6
tunnel mpls traffic-eng autoroute announce
tunnel mpls traffic-eng priority 6 6
tunnel mpls traffic-eng path-option 10 segment-routing
tunnel mpls traffic-eng path-selection metric igp
tunnel mpls traffic-eng 10
```

## Verifying Configuration of the SR-TE Tunnels

Use the **show mpls traffic-eng tunnels *tunnel-number*** command to verify the configuration of the SR-TE tunnels.

## Verifying Tunnel 1

```

Name: R1_t1                                     (Tunnell) Destination: 6.6.6.6
Status:
  Admin: up          Oper: up          Path: valid          Signalling: connected
  path option 10, (SEGMENT-ROUTING) type explicit IP_PATH (Basis for Setup)
Config Parameters:
  Bandwidth: 0          kbps (Global) Priority: 6 6  Affinity: 0x0/0xFFFF
  Metric Type: IGP (interface)
  Path Selection:
    Protection: any (default)
  Path-invalidation timeout: 45000 msec (default), Action: Tear
  AutoRoute: enabled LockDown: disabled Loadshare: 10 [200000000]
  auto-bw: disabled
  Fault-OAM: disabled, Wrap-Protection: disabled, Wrap-Capable: No
Active Path Option Parameters:
  State: explicit path option 10 is active
  BandwidthOverride: disabled LockDown: disabled Verbatim: disabled
History:
  Tunnel:
    Time since created: 6 days, 19 hours
    Time since path change: 2 seconds
    Number of LSP IDs (Tun_Instances) used: 1814
  Current LSP: [ID: 1814]
    Uptime: 2 seconds
    Selection: reoptimization
  Prior LSP: [ID: 1813]
    ID: path option unknown
    Removal Trigger: configuration changed
Tun_Instance: 1814
Segment-Routing Path Info (isis level-1)
  Segment0[Node]: 4.4.4.4, Label: 114
  Segment1[Node]: 5.5.5.5, Label: 115
  Segment2[Node]: 6.6.6.6, Label: 116

```

## Verifying Tunnel 2

```

Name: R1_t2                                     (Tunnell) Destination: 6.6.6.6
Status:
  Admin: up          Oper: up          Path: valid          Signalling: connected
  path option 10, (SEGMENT-ROUTING) type explicit IA_PATH (Basis for Setup)
Config Parameters:
  Bandwidth: 0          kbps (Global) Priority: 6 6  Affinity: 0x0/0xFFFF
  Metric Type: IGP (interface)
  Path Selection:
    Protection: any (default)
  Path-invalidation timeout: 45000 msec (default), Action: Tear
  AutoRoute: enabled LockDown: disabled Loadshare: 10 [200000000]
  auto-bw: disabled
  Fault-OAM: disabled, Wrap-Protection: disabled, Wrap-Capable: No
Active Path Option Parameters:
  State: explicit path option 10 is active
  BandwidthOverride: disabled LockDown: disabled Verbatim: disabled
History:
  Tunnel:
    Time since created: 6 days, 19 hours, 1 minutes
    Time since path change: 1 seconds
    Number of LSP IDs (Tun_Instances) used: 1815
  Current LSP: [ID: 1815]
    Uptime: 1 seconds

```

```

Prior LSP: [ID: 1814]
  ID: path option unknown
  Removal Trigger: configuration changed
Tun_Instance: 1815
Segment-Routing Path Info (isis level-1)
Segment0[ - ]: Label: 114
Segment1[ - ]: Label: 115
Segment2[ - ]: Label: 116

```

## Verifying Tunnel 3

```

Name: R1_t3                                     (Tunnel) Destination: 6.6.6.6
Status:
  Admin: up           Oper: up           Path: valid           Signalling: connected
  path option 10, (SEGMENT-ROUTING) type explicit MIXED_PATH (Basis for Setup)
Config Parameters:
  Bandwidth: 0          kbps (Global)  Priority: 6 6  Affinity: 0x0/0xFFFF
  Metric Type: IGP (interface)
  Path Selection:
    Protection: any (default)
  Path-invalidation timeout: 45000 msec (default), Action: Tear
  AutoRoute: enabled  LockDown: disabled Loadshare: 10 [200000000]
  auto-bw: disabled
  Fault-OAM: disabled, Wrap-Protection: disabled, Wrap-Capable: No
Active Path Option Parameters:
  State: explicit path option 10 is active
  BandwidthOverride: disabled  LockDown: disabled  Verbatim: disabled
History:
  Tunnel:
    Time since created: 6 days, 19 hours, 2 minutes
    Time since path change: 2 seconds
    Number of LSP IDs (Tun_Instances) used: 1816
  Current LSP: [ID: 1816]
    Uptime: 2 seconds
    Selection: reoptimization
  Prior LSP: [ID: 1815]
    ID: path option unknown
    Removal Trigger: configuration changed
  Tun_Instance: 1816
Segment-Routing Path Info (isis level-1)
Segment0[Node]: 2.2.2.2, Label: 112
Segment1[Node]: 3.3.3.3, Label: 113
Segment2[ - ]: Label: 115
Segment3[ - ]: Label: 116

```

## Verifying Tunnel 4

```

Name: R1_t4                                     (Tunnel) Destination: 6.6.6.6
Status:
  Admin: up           Oper: up           Path: valid           Signalling: connected
  path option 10, (SEGMENT-ROUTING) type dynamic (Basis for Setup, path weight 30)
Config Parameters:
  Bandwidth: 0          kbps (Global)  Priority: 6 6  Affinity: 0x0/0xFFFF
  Metric Type: IGP (interface)
  Path Selection:
    Protection: any (default)
  Path-invalidation timeout: 45000 msec (default), Action: Tear
  AutoRoute: enabled  LockDown: disabled Loadshare: 10 [200000000]
  auto-bw: disabled

```

```

    Fault-OAM: disabled, Wrap-Protection: disabled, Wrap-Capable: No
Active Path Option Parameters:
    State: dynamic path option 10 is active
    BandwidthOverride: disabled LockDown: disabled Verbatim: disabled
History:
    Tunnel:
        Time since created: 6 days, 19 hours
        Time since path change: 2 seconds
        Number of LSP IDs (Tun_Instances) used: 1813
    Current LSP: [ID: 1813]
        Uptime: 2 seconds
    Prior LSP: [ID: 1806]
        ID: path option unknown
        Removal Trigger: configuration changed
Tun_Instance: 1813
Segment-Routing Path Info (isis level-1)
    Segment0[Link]: 192.168.2.1 - 192.168.2.2, Label: 17
    Segment1[Link]: 192.168.4.2 - 192.168.4.1, Label: 25
    Segment2[Link]: 192.168.8.1 - 192.168.8.2, Label: 300

```

## Verifying Tunnel 5

```

Name: R1_t5                                     (Tunnell) Destination: 6.6.6.6
Status:
    Admin: up          Oper: up          Path: valid          Signalling: connected
    path option 10, type segment-routing (Basis for Setup)
Config Parameters:
    Bandwidth: 0          kbps (Global) Priority: 6 6 Affinity: 0x0/0xFFFF
    Metric Type: IGP (interface)
    Path Selection:
        Protection: any (default)
    Path-invalidation timeout: 45000 msec (default), Action: Tear
    AutoRoute: enabled LockDown: disabled Loadshare: 10 [200000000]
    auto-bw: disabled
    Fault-OAM: disabled, Wrap-Protection: disabled, Wrap-Capable: No
Active Path Option Parameters:
    State: segment-routing path option 10 is active
    BandwidthOverride: disabled LockDown: disabled Verbatim: disabled
History:
    Tunnel:
        Time since created: 6 days, 19 hours, 4 minutes
        Time since path change: 14 seconds
        Number of LSP IDs (Tun_Instances) used: 1817
    Current LSP: [ID: 1817]
        Uptime: 14 seconds
        Selection: reoptimization
    Prior LSP: [ID: 1816]
        ID: path option unknown
        Removal Trigger: configuration changed
Tun_Instance: 1817
Segment-Routing Path Info (isis level-1)
    Segment0[Node]: 6.6.6.6, Label: 116

```

## SR-TE LSP Instantiation

A Traffic Engineered (TE) tunnel is a container of one or more instantiated TE LSPs. An SR-TE LSP is instantiated by configuring 'segment-routing' on the path-option of the TE tunnel. The traffic mapped to the tunnel is forwarded over the primary SR-TE instantiated LSP.

Multiple path-options can also be configured under the same tunnel. Each path-option is assigned a preference index or a path-option index that is used to determine the more favorable path-option for instantiating the primary LSP—the lower the path-option preference index, the more favorable the path-option. The other less favorable path-options under the same TE tunnel are considered secondary path-options and may be used once the currently used path-option is invalidated (for example, due to a failure on the path).

**Note**

A forwarding state is maintained for the primary LSP only.

## SR-TE LSP Explicit Null

MPLS-TE tunnel headend does not impose explicit-null at the bottom of the stack. When penultimate hop popping (PHP) is enabled for SR prefix SIDs or when an adjacency SID is the last hop of the SR-TE LSP, the packet may arrive at the tailend without a transport label. However, sometimes, it is desirable that the packet arrive at the tailend with explicit-null label, and in such cases, the headend imposes an explicit-null label at the top of the label stack.

## SR-TE LSP Path Verification

SR-TE tunnel functionality requires that the headend perform initial verification of the tunnel path as well as the subsequent tracking of the reachability of the tunnel tailend and traversed segments.

Path verification for SR-TE LSP paths is triggered whenever MPLS-TE is notified of any topology changes or SR SID updates.

The SR-TE LSP validation steps consist of the following checks:

### Topology Path Validation

The headend validates the path of an SR-TE LSP for connectivity against the TE topology. MPLS-TE headend checks if links corresponding to the adjacency SIDs are connected in the TE topology.

For newly instantiated SR-TE LSPs, if the headend detects a discontinuity on any link of the SR-TE path, that path is considered invalid and is not used. If the tunnel has other path-options with valid paths, those paths are used to instantiate the tunnel LSP.

For TE tunnels with existing instantiated SR-TE LSP, if the headend detects a discontinuity on any link, the headend assumes that a fault has occurred on that link. In this case, the local repair protection, such as the IP FRR, comes in to effect. The IGPs continue to sustain the protected adjacency label and associated forwarding after the adjacency is lost for some time. This allows the head-ends enough time to reroute the tunnels onto different paths that are not affected by the same failure. The headend starts a tunnel invalidation timer once it detects the link failure to attempt to reroute the tunnel onto other available path-options with valid paths.

If the TE tunnel is configured with other path-options that are not affected by the failure and are validated, the headend uses one of those path-options to reroute (and re-optimize) the tunnel by instantiating a new primary LSP for the tunnel using the unaffected path.

If no other valid path-options exist under the same tunnel, or if the TE tunnel is configured with only one path-option that is affected by the failure, the headend starts an invalidation timer after which it brings the tunnel state to 'down'. This action avoids black-holing the traffic flowing over the affected SR-TE LSP, and allows services riding over the tunnel to reroute over different available paths at the headend. There is an

invalidation drop configuration that keeps the tunnel 'up', but drops the traffic when the invalidation timer expires.

For intra-area SR-TE LSPs, the headend has full visibility over the LSP path, and validates the path to the ultimate LSP destination. However, for interarea LSPs, the headend has partial visibility over the LSP path—only up to the first ABR. In this case, the headend can only validate the path from the ingress to the first ABR. Any failure along the LSP beyond the first ABR node is invisible to the headend, and other mechanisms to detect such failures, such as BFD over LSP are assumed.

## SR SID Validation

SID hops of an SR-TE LSP are used to determine the outgoing MPLS label stack to be imposed on the outgoing packets carried over the SR-TE LSP of a TE tunnel. A database of global and local adjacency-SIDs is populated from the information received from IGP and maintained in MPLS-TE. Using a SID that is not available in the MPLS TE database invalidates the path-option using the explicit-path. The path-option, in this case, is not used to instantiate the SR TE LSP. Also, withdrawing, adding, or modifying a SID in the MPLS-TE SID-database, results in the MPLS-TE headend verifying all tunnels with SR path-options (in-use or secondary) and invokes proper handling.

## LSP Egress Interface

When the SR-TE LSP uses an adjacency-SID for the first path hop, TE monitors the interface state and IGP adjacency state associated with the adjacency-SID and the node that the SR-TE LSP egresses on. If the interface or adjacency goes down, TE can assume that a fault occurred on the SR-TE LSP path and take the same reactive actions described in the previous sections.



**Note** When the SR-TE LSP uses a prefix-SID for the first hop, TE cannot directly infer on which interface the tunnel egresses. TE relies on the IP reachability information of the prefix to determine if connectivity to the first hop is maintained.

## IP Reachability Validation

MPLS-TE validates that the nodes corresponding to the prefix-SIDs are IP reachable before declaring the SR path valid. MPLS-TE detects path changes for the IP prefixes corresponding to the adjacency or prefix SIDs of the SR-TE LSP path. If the node announcing a specific SID loses IP reachability due to a link or node failure, MPLS-TE is notified of the path change (no path). MPLS-TE reacts by invalidating the current SR-TE LSP path, and may use other path-options with a valid path, if any to instantiate a new SR-TE LSP.



**Note** Since IP-FRR does not offer protection against failure of a node that is being traversed by an SR-TE LSP (such as, a prefix-SID failure along the SR-TE LSP path), the headend immediately reacts to IP route reachability loss for prefix-SID node by setting the tunnel state to 'down' and removes the tunnel forwarding entry if there are no other path-options with valid path for the affected tunnel.

## Tunnel Path Affinity Validation

The affinity of a tunnel path can be specified using the command **tunnel mpls traffic-eng affinity** under the tunnel interface.

The headend validates that the specified SR path is compliant with the configured affinity. This requires that the paths of each segment of the SR path be validated against the specified constraint. The path is declared invalid against the configured affinity constraints if at least a single segment of the path does not satisfy the configured affinity.

```
interface Tunnell
no ip address
tunnel mode mpls traffic-eng
tunnel destination 5.5.5.5
tunnel mpls traffic-eng priority 5 5
tunnel mpls traffic-eng bandwidth 100
tunnel mpls traffic-eng affinity 0x1 mask 0xFFFF
tunnel mpls traffic-eng path-option 10 dynamic segment-routing
Router# show tunnel ??
Name: R1_t1                                     (Tunnell) Destination: 5.5.5.5
Status:
  Admin: up          Oper: up          Path: valid          Signalling: connected
  path option 10, (SEGMENT-ROUTING) type dynamic (Basis for Setup, path weight 20)
Config Parameters:
  Bandwidth: 100          kbps (Global)  Priority: 5 5    Affinity: 0x1/0xFFFF
  Metric Type: TE (default)
  Path Selection:
    Protection: any (default)
  Path-selection Tiebreaker:
    Global: not set  Tunnel Specific: not set  Effective: min-fill (default)
  Hop Limit: disabled
  Cost Limit: disabled
  Path-invalidation timeout: 10000 msec (default), Action: Tear
  AutoRoute: disabled LockDown: disabled Loadshare: 100 [0] bw-based
  auto-bw: disabled
  Fault-OAM: disabled, Wrap-Protection: disabled, Wrap-Capable: No
Active Path Option Parameters:
  State: dynamic path option 10 is active
  BandwidthOverride: disabled LockDown: disabled Verbatim: disabled
Node Hop Count: 2
History:
  Tunnel:
    Time since created: 10 minutes, 54 seconds
    Time since path change: 34 seconds
    Number of LSP IDs (Tun_Instances) used: 55
  Current LSP: [ID: 55]
    Uptime: 34 seconds
  Prior LSP: [ID: 49]
    ID: path option unknown
    Removal Trigger: tunnel shutdown
Tun_Instance: 55
Segment-Routing Path Info (isis level-1)
  Segment0[Link]: 192.168.2.1 - 192.168.2.2, Label: 46
  Segment1[Link]: 192.168.4.2 - 192.168.4.1, Label: 49
```

## Configuring Affinity on an Interface

```
interface GigabitEthernet2
ip address 192.168.2.1 255.255.255.0
ip router isis 1
negotiation auto
mpls traffic-eng tunnels
mpls traffic-eng attribute-flags 0x1
isis network point-to-point
ip rsvp bandwidth
```



## Tunnel Path Resource Avoidance Validation

You can specify a set of addresses to be validated as excluded from being traversed by SR-TE tunnel packets. To achieve this, the headend runs the per-segment verification checks and validates that the specified node, prefix or link addresses are indeed excluded from the tunnel in the SR path. The tunnel resource avoidance checks can be enabled per path using the following commands. The list of addresses to be excluded are defined and the name of the list is referenced in the path-option.

```
interface tunnel100
 tunnel mpls traffic-eng path-option 1 explicit name EXCLUDE segment-routing
 ip explicit-path name EXCLUDE enable
  exclude-address 192.168.0.2
  exclude-address 192.168.0.4
  exclude-address 192.168.0.3
!
```

## Tunnel Path Loop Validation

The SR path is a concatenation of SR segments (combination of prefix and adjacency SIDs). It is possible that any of the traversed segment's underlying paths may traverse through the ingress of the tunnel. In this case, packets that are mapped on the SR tunnel may loop back again to the headend. To avoid this sub-optimal path, the headend detects and invalidates a looping SR path through the ingress node.

Loop path validation is implicitly enabled on SR path. However, it is possible to disable this validation by using the **verbatim** path-option keyword associated with the tunnel path-option.

The following is an example of the **verbatim** path-option keyword when IP address 6.6.6.6 is in a different area:

```
interface Tunnell
 tunnel mode mpls traffic-eng
 tunnel destination 6.6.6.6
 tunnel mpls traffic-eng priority 5 5
 tunnel mpls traffic-eng bandwidth 100
 tunnel mpls traffic-eng path-option 10 explicit name NODE_PATH segment-routing verbatim
Name: RI_tl1 (Tunnell) Destination: 6.6.6.6
Status:
  Admin: up          Oper: up          Path: valid          Signalling: connected
  path option 10, (SEGMENT-ROUTING) type explicit (verbatim) NODE_PATH (Basis for Setup)
Config Parameters:
  Bandwidth: 100      kbps (Global)  Priority: 5 5      Affinity: 0x0/0xFFFF
  Metric Type: TE (default)
  Path Selection:
    Protection: any (default)
  Path-selection Tiebreaker:
    Global: not set   Tunnel Specific: not set   Effective: min-fill (default)
  Hop Limit: disabled [ignore: Verbatim Path Option]
  Cost Limit: disabled
  Path-invalidation timeout: 10000 msec (default), Action: Tear
  AutoRoute: disabled LockDown: disabled Loadshare: 100 [0] bw-based
  auto-bw: disabled
  Fault-OAM: disabled, Wrap-Protection: disabled, Wrap-Capable: No
Active Path Option Parameters:
  State: explicit path option 10 is active
  BandwidthOverride: disabled LockDown: disabled Verbatim: enabled
History:
  Tunnel:
    Time since created: 7 minutes, 43 seconds
    Time since path change: 0 seconds
    Number of LSP IDs (Tun_Instances) used: 49
```

```

Current LSP: [ID: 49]
  Uptime: 0 seconds
Prior LSP: [ID: 48]
  ID: path option unknown
  Removal Trigger: signalling shutdown
Tun_Instance: 49
Segment-Routing Path Info (isis level-1)
  Segment0[Node]: 2.2.2.2, Label: 20012
  Segment1[Node]: 3.3.3.3, Label: 20013
  Segment2[ - ]: Label: 20016

```

## SR-TE Traffic Load Balancing

SR-TE tunnels support the following load-balancing options:

### Load Balancing on Port Channel TE Links

Port Channel interfaces carry the SR-TE LSP traffic. This traffic load balances over port channel member links as well as over bundle interfaces on the head or mid of an SR-TE LSP.

### Load Balancing on ECMPs

While using the equal cost multi path protocol (ECMP), the path to a specific prefix-SID may point to multiple next-hops. And if the SR-TE LSP path traverses one or more prefix-SIDs that have ECMP, the SR-TE LSP traffic load-balances on the ECMP paths of each traversed prefix-SID from any midpoint traversed node along the SR-TE LSP path.



**Note**

ECMP within a single SR-TE tunnel is not supported.

### Load Balancing on Multiple Tunnels

Effective Cisco IOS XE Bengaluru 17.5.1, ECMP across multiple SR-TE tunnels is supported on RSP3. See [ECMP Over SR-TE, on page 221](#) for more details.

## SR-TE Tunnel Re-optimization

TE tunnel re-optimization occurs when the headend determines that there is a more optimal path available than the one currently used. For example, if there is a failure along the SR-TE LSP path, the headend could detect and revert to a more optimal path by triggering re-optimization.

Tunnels that instantiate SR-TE LSP can re-optimize without affecting the traffic carried over the tunnel.

Re-optimization can occur because:

- The explicit path hops used by the primary SR-TE LSP explicit path are modified,
- The headend determines the currently used path-option are invalid due to either a topology path disconnect, or a missing SID in the SID database that is specified in the explicit-path
- A more favorable path-option (lower index) becomes available

When the headend detects a failure on a protected SR adjacency-SID that is traversed by an SR-TE LSP, it starts the invalidation timer. If the timer expires and the headend is still using the failed path because it is

unable to reroute on a different path, the tunnel state is brought 'down' to avoid black-holing the traffic. Once the tunnel is down, services on the tunnel converge to take a different path.

The following is a sample output of a manual re-optimization example. In this example, the path-option is changed from '10' to '20'.

```
Router# mpls traffic-eng reoptimize tunnel 1 path-option 20
The targeted path-option is not in lock down mode. Continue? [no]: yes
Router# show mpls traffic-eng tunnels tunnell
Name: Rl_tl1 (Tunnell) Destination: 6.6.6.6
Status:
  Admin: up      Oper: up      Path: valid      Signalling: connected
  path option 20, (SEGMENT-ROUTING) type explicit IP_PATH (Basis for Setup)
  path option 10, (SEGMENT-ROUTING) type dynamic
Config Parameters:
  Bandwidth: 0      kbps (Global) Priority: 6 6 Affinity: 0x0/0xFFFF
  Metric Type: IGP (interface)
  Path Selection:
    Protection: any (default)
  Path-invalidation timeout: 45000 msec (default), Action: Tear
  AutoRoute: enabled LockDown: disabled Loadshare: 10 [200000000]
  auto-bw: disabled
  Fault-OAM: disabled, Wrap-Protection: disabled, Wrap-Capable: No
Active Path Option Parameters:
  State: explicit path option 20 is active
  BandwidthOverride: disabled LockDown: disabled Verbatim: disabled
History:
  Tunnel:
    Time since created: 6 days, 19 hours, 9 minutes
    Time since path change: 14 seconds
    Number of LSP IDs (Tun_Instances) used: 1819
  Current LSP: [ID: 1819]
    Uptime: 17 seconds
    Selection: reoptimization
  Prior LSP: [ID: 1818]
    ID: path option unknown
    Removal Trigger: reoptimization completed
Tun_Instance: 1819
Segment-Routing Path Info (isis level-1)
  Segment0[Node]: 4.4.4.4, Label: 114
  Segment1[Node]: 5.5.5.5, Label: 115
  Segment2[Node]: 6.6.6.6, Label: 116
```




---

**Note** SR-TE does not support lossless re-optimization with multiple path options.

---




---

**Note** When FRR is configured and the primary path is brought back up, re-optimization time is in the order of seconds due to microloop.

---

## SR-TE With lockdown Option

The **lockdown** option only prevents SR-TE from re-optimizing to a better path. However, it does not prevent signaling the existence of a new path.

```
interface Tunnell
```

```

ip unnumbered Loopback1 poll point-to-point
tunnel mode mpls traffic-eng
tunnel destination 6.6.6.6
tunnel mpls traffic-eng autoroute announce
tunnel mpls traffic-eng priority 6 6
tunnel mpls traffic-eng path-option 10 segment-routing lockdown
tunnel mpls traffic-eng path-option 20 segment-routing
tunnel mpls traffic-eng path-selection metric igp
tunnel mpls traffic-eng 10
Router# show mpls traffic-eng tunnels tunnel1
Name: csr551_t1 (Tunnel1) Destination: 6.6.6.6
Status:
  Admin: up      Oper: up      Path: valid      Signalling: connected
  path option 10, (LOCKDOWN) type segment-routing (Basis for Setup)
Config Parameters:
  Bandwidth: 0      kbps (Global) Priority: 6 6      Affinity: 0x0/0xFFFF
  Metric Type: IGP (interface)
  Path Selection:
    Protection: any (default)
  Path-invalidation timeout: 45000 msec (default), Action: Tear
  AutoRoute: enabled LockDown: enabled Loadshare: 10 [200000000]
  auto-bw: disabled
  Fault-OAM: disabled, Wrap-Protection: disabled, Wrap-Capable: No
Active Path Option Parameters:
  State: segment-routing path option 10 is active
  BandwidthOverride: disabled LockDown: enabled Verbatim: disabled
History:
  Tunnel:
    Time since created: 6 days, 19 hours, 22 minutes
    Time since path change: 1 minutes, 26 seconds
    Number of LSP IDs (Tun_Instances) used: 1822
  Current LSP: [ID: 1822]
    Uptime: 1 minutes, 26 seconds
    Selection: reoptimization
  Prior LSP: [ID: 1821]
    ID: path option unknown
    Removal Trigger: configuration changed
  Tun_Instance: 1822
  Segment-Routing Path Info (isis level-1)
    Segment0[Node]: 6.6.6.6, Label: 116

```

## SR-TE Tunnel Protection

Protection for SR TE tunnels can take any of the following alternatives:



### Note

50-millisecond traffic protection is not guaranteed for path protection scenarios.

## IP-FRR Local Repair Protection

On an SR-TE LSP headend or mid-point node, IP-FRR is used to compute and program the backup protection path for the prefix-SID or adjacency-SID label.

With IP-FRR, backup repair paths are pre-computed and pre-programmed by IGP *before* a link or node failure. The failure of a link triggers its immediate withdrawal from the TE topology (link advertisement withdrawal). This allows the headend to detect the failure of an SR-TE LSP traversing the failed adjacency-SID.

When a protected adjacency-SID fails, the failed adjacency-SID label and associated forwarding are kept functional for a specified period of time (5 to 15 minutes) to allow all SR TE tunnel head-ends to detect and react to the failure. Traffic using the adjacency-SID label continues to be FRR protected even if there are subsequent topology updates that change the backup repair path. In this case, the IGP updates the backup repair path while FRR is active to reroute traffic on the newly-computed backup path.

When the primary path of a protected prefix-SID fails, the PLR reroutes to the backup path. The headend remains transparent to the failure and continues to use the SR-TE LSP as a valid path.

IP-FRR provides protection for adjacency and prefix-SIDs against link failures only.

## Tunnel Path Protection

Path protection is the instantiation of one or more standby LSPs to protect against the failure of the primary LSP of a single TE tunnel.

Path protection protects against failures by pre-computing and pre-provisioning secondary paths that are failure diverse with the primary path-option under the same tunnel. This protection is achieved by computing a path that excludes prefix-SIDs and adjacency-SIDs traversed by the primary LSP or by computing a path that excludes SRLGs of the primary SR-TE LSP path.

If the primary SR-TE LSP fails, at least one standby SR-TE LSP is used for the tunnel. Multiple secondary path-options can be configured to be used as standby SR-TE LSPs paths.

## SR-TE and TI-LFA

### Restrictions for Using SR-TE and TI-LFA

- In case of primary and secondary path switchover, a microloop is created between routers. Use the **microloop avoidance rib-update-delay** command to bring down the convergence time

Consider the following topology:

```

-----ixia-2
|
------(R4)-----
||
||
(R3) (R1) ----ixia-1
||
||
------(R2)-----

router isis ipfrr
net 49.0001.0120.1201.2012.00
is-type level-2-only
ispf level-2
metric-style wide
log-adjacency-changes

```

```

segment-routing mpls
segment-routing prefix-sid-map advertise-local
fast-reroute per-prefix level-2 all
fast-reroute ti-lfa level-2
microloop avoidance rib-update-delay 10000

```

- Before you configure SR-TE for TI-LFA, you must enable TI-LFA is enabled on all nodes. For more information, see *Using Segment Routing with OSPF*.

```

mpls traffic-eng tunnels
!
segment-routing mpls
connected-prefix-sid-map
  address-family ipv4
    1.1.1.1/32 index 11 range 1
  exit-address-family
!
interface Loopback1
ip address 1.1.1.1 255.255.255.255
ip router isis 1
!
interface Tunnel1
ip unnumbered Loopback1 poll point-to-point
tunnel mode mpls traffic-eng
tunnel destination 6.6.6.6
tunnel mpls traffic-eng autoroute announce
tunnel mpls traffic-eng path-option 10 explicit name IP_PATH segment-routing
!
interface GigabitEthernet2
ip address 192.168.1.1 255.255.255.0
ip router isis 1
negotiation auto
mpls traffic-eng tunnels
isis network point-to-point
!
interface GigabitEthernet3
ip address 192.168.2.1 255.255.255.0
ip router isis 1
negotiation auto
mpls traffic-eng tunnels
isis network point-to-point
!
router isis 1
net 49.0001.0010.0100.1001.00
is-type level-1
ispf level-1
metric-style wide
log-adjacency-changes
segment-routing mpls
fast-reroute per-prefix level-1 all
fast-reroute ti-lfa level-1
mpls traffic-eng router-id Loopback1
mpls traffic-eng level-1
!
ip explicit-path name IP_PATH enable
next-address 4.4.4.4
next-address 5.5.5.5
next-address 6.6.6.6

```

- To reduce or minimize traffic loss after a high availability (HA) switchover, MPLS TE NSR and IS-IS NSF must be enabled.

Use the **mpls traffic-eng nsr** command in global EXEC mode.

```
mpls traffic-eng nsr
```

Use the **nsf** command under IS-IS or OSPF.

```
router isis
nsf cisco
nsf interval 0
```

- The Cisco ASR routers support 500 SR-TE tunnels with two transport labels, two TI-LFA protection labels and one service label.
- SSO is not supported with SR-TE on the Cisco RSP2 Module.
- For TI-LFA restrictions, see [Restrictions for the TI-LFA](#).

## Verifying the SR-TE With TI\_LFA Configuration

```
Router# show mpls traffic-eng tunnels tunnel1
Name: PE1 (Tunnel1) Destination: 6.6.6.6
Status:
  Admin: up      Oper: up      Path: valid      Signalling: connected
  path option 10, (SEGMENT-ROUTING) type explicit IP_PATH (Basis for Setup)
Config Parameters:
  Bandwidth: 0      kbps (Global) Priority: 7 7 Affinity: 0x0/0xFFFF
  Metric Type: TE (default)
  Path Selection:
    Protection: any (default)
  Path-invalidation timeout: 45000 msec (default), Action: Tear
  AutoRoute: enabled LockDown: disabled Loadshare: 0 [0] bw-based
  auto-bw: disabled
  Fault-OAM: disabled, Wrap-Protection: disabled, Wrap-Capable: No
Active Path Option Parameters:
  State: explicit path option 10 is active
  BandwidthOverride: disabled LockDown: disabled Verbatim: disabled
History:
  Tunnel:
    Time since created: 4 hours, 25 minutes
    Time since path change: 4 hours, 21 minutes
    Number of LSP IDs (Tun_Instances) used: 37
    Current LSP: [ID: 37]
    Uptime: 4 hours, 21 minutes
  Tun_Instance: 37
  Segment-Routing Path Info (isis level-1)
    Segment0[Node]: 4.4.4.4, Label: 16014
    Segment1[Node]: 5.5.5.5, Label: 16015
    Segment2[Node]: 6.6.6.6, Label: 16016
Router# show isis fast-reroute ti-lfa tunnel

Tag 1:
Fast-Reroute TI-LFA Tunnels:
Tunnel Interface Next Hop      End Point      Label      End Point Host
MP1     Gi2       192.168.1.2     6.6.6.6       16016      SR_R6
MP2     Gi3       192.168.2.2     6.6.6.6       16016      SR_R6
Router# show frr-manager client
client-name
  ISIS interfaces detail
TunnelI/F : MP1
Type : SR
Next-hop : 192.168.1.2
End-point : 6.6.6.6
OutI/F : Gi2
```

```

Adjacency State : 1
Prefix0 : 6.6.6.6(Label : 16016)
TunnelI/F : MP2
Type : SR
Next-hop : 192.168.2.2
End-point : 6.6.6.6
OutI/F : Gi3
Adjacency State : 1
Prefix0 : 6.6.6.6(Label : 16016)
Router# show ip cef 6.6.6.6 internal

6.6.6.6/32, epoch 2, RIB[I], refcnt 6, per-destination sharing
sources: RIB, LTE
feature space:
  IPRM: 0x00028000
  Broker: linked, distributed at 1st priority
  LFD: 6.6.6.6/32 1 local label
  sr local label info: global/16016 [0x1A]
    contains path extension list
    sr disposition chain 0x7FC6B0BF2AF0
      label implicit-null
      IP midchain out of Tunnel1
      label 16016
      FRR Primary
        <primary: label 16015
          TAG adj out of GigabitEthernet3, addr 192.168.2.2>
    sr label switch chain 0x7FC6B0BF2B88
      label implicit-null
      TAG midchain out of Tunnel1
      label 16016
      FRR Primary
        <primary: label 16015
          TAG adj out of GigabitEthernet3, addr 192.168.2.2>
  ifnums:
    Tunnel1(13)
  path list 7FC6B0BBDDE0, 3 locks, per-destination, flags 0x49 [shble, rif, hwc]
  path 7FC7144D4300, share 1/1, type attached nexthop, for IPv4
    MPLS short path extensions: [rib | prfmfi | lblmrg | srlbl] MOI flags = 0x3 label
implicit-null
  nexthop 6.6.6.6 Tunnel1, IP midchain out of Tunnel1 7FC6B0BBB440
output chain:
  IP midchain out of Tunnel1 7FC6B0BBB440
  label [16016|16016]
  FRR Primary (0x7FC714515460)
    <primary: label 16015
      TAG adj out of GigabitEthernet3, addr 192.168.2.2 7FC6B0BBB630>
    <repair: label 16015
      label 16014
      TAG midchain out of MPLS-SR-Tunnel1 7FC6B0BBAA90
      label 16016
      TAG adj out of GigabitEthernet2, addr 192.168.1.2 7FC6B0BBBA10>

```

**Note**

To ensure a less than 50-msec traffic protection with TI-LFA, SR-TE with dynamic path option must use the backup adjacency SID.

To create an SR-TE with dynamic path option, use the following configuration on every router in the topology:

```

router isis 1
fast-reroute per-prefix level-1 all

```



At the tunnel headend router:

```
interface Tunnel1
ip unnumbered Loopback1 poll point-to-point
tunnel mode mpls traffic-eng
tunnel destination 6.6.6.6
tunnel mpls traffic-eng autoroute announce
tunnel mpls traffic-eng path-option 1 dynamic segment-routing
tunnel mpls traffic-eng path-selection segment-routing adjacency protected
```

## Configuring TI-LFA With Mapping Server

Consider the following topology:

- IXIA-2 injects IS-IS prefixes, and IXIA-1 sends one-way traffic to IXIA-2
- In R1 10,000 prefixes are configured in the segment-routing mapping-server

The configuration on R1 is:

```
conf t
segment-routing mpls
global-block 16 20016
!
connected-prefix-sid-map
address-family ipv4
11.11.11.11/32 index 11 range 1
exit-address-family
!
!
mapping-server
!
prefix-sid-map
address-family ipv4
120.0.0.0/24 index 2 range 1 attach
200.0.0.0/24 index 1 range 1 attach
192.168.0.0/24 index 100 range 10000 attach
exit-address-family
!
!
!
!
interface Loopback0
ip address 11.11.11.11 255.255.255.255
ip router isis ipfrr
!
interface GigabitEthernet0/1/0
ip address 14.0.0.1 255.255.255.0
ip router isis ipfrr
negotiation auto
isis network point-to-point
!
interface GigabitEthernet0/1/2
ip address 11.0.0.1 255.255.255.0
ip router isis ipfrr
negotiation auto
isis network point-to-point
!
interface GigabitEthernet0/1/4
ip address 200.0.0.1 255.255.255.0
ip router isis ipfrr
negotiation auto
```

```

isis network point-to-point
!
router isis ipfrr
net 49.0001.0110.1101.1011.00
is-type level-2-only
metric-style wide
log-adjacency-changes
nsf cisco
segment-routing mpls
segment-routing prefix-sid-map advertise-local
fast-reroute per-prefix level-2 all
fast-reroute ti-lfa level-2
microloop avoidance rib-update-delay 10000

```

On R2 the configuration is

```

conf t
!
!
segment-routing mpls
!
connected-prefix-sid-map
address-family ipv4
12.12.12.12/32 index 12 range 1
exit-address-family
!
!
interface Loopback0
ip address 12.12.12.12 255.255.255.255
ip router isis ipfrr
!
interface GigabitEthernet0/1/0
ip address 12.0.0.1 255.255.255.0
ip router isis ipfrr
negotiation auto
isis network point-to-point
!
interface GigabitEthernet0/1/1
ip address 11.0.0.2 255.255.255.0
ip router isis ipfrr
negotiation auto
isis network point-to-point
!
router isis ipfrr
net 49.0001.0120.1201.2012.00
is-type level-2-only
metric-style wide
log-adjacency-changes
nsf cisco
segment-routing mpls
segment-routing prefix-sid-map advertise-local
fast-reroute per-prefix level-2 all
fast-reroute ti-lfa level-2
microloop avoidance rib-update-delay 10000
!

```

On R3 the configuration is

```

conf t
!
mpls traffic-eng tunnels
!
segment-routing mpls

```

```
!  
connected-prefix-sid-map  
address-family ipv4  
13.13.13.13/32 index 13 range 1  
exit-address-family  
!  
!  
interface Loopback0  
ip address 13.13.13.13 255.255.255.255  
ip router isis ipfrr  
!  
interface GigabitEthernet0/0/4  
ip address 13.0.0.1 255.255.255.0  
ip router isis ipfrr  
load-interval 30  
speed 1000  
no negotiation auto  
isis network point-to-point  
!  
interface GigabitEthernet0/0/5  
ip address 12.0.0.2 255.255.255.0  
ip router isis ipfrr  
negotiation auto  
isis network point-to-point  
!  
router isis ipfrr  
net 49.0001.0130.1301.3013.00  
is-type level-2-only  
metric-style wide  
log-adjacency-changes  
nsf cisco  
segment-routing mpls  
segment-routing prefix-sid-map advertise-local  
fast-reroute per-prefix level-2 all  
fast-reroute ti-lfa level-2  
microloop avoidance rib-update-delay 10000  
!
```

On R4 the configuration is:

```
conf t  
!  
mpls traffic-eng tunnels  
!  
segment-routing mpls  
!  
connected-prefix-sid-map  
address-family ipv4  
14.14.14.14/32 index 14 range 1  
exit-address-family  
!  
!  
interface Loopback0  
ip address 14.14.14.14 255.255.255.255  
ip router isis ipfrr  
!  
interface GigabitEthernet0/0/0  
ip address 14.0.0.2 255.255.255.0  
ip router isis ipfrr  
negotiation auto  
isis network point-to-point  
!  
interface GigabitEthernet0/0/3  
ip address 13.0.0.2 255.255.255.0
```

```
ip router isis ipfrr
speed 1000
no negotiation auto
isis network point-to-point
!
interface GigabitEthernet0/0/5
ip address 120.0.0.1 255.255.255.0
ip router isis ipfrr
speed 1000
no negotiation auto
isis network point-to-point
!
router isis ipfrr
net 49.0001.0140.1401.4014.00
is-type level-2-only
metric-style wide
log-adjacency-changes
nsf cisco
segment-routing mpls
segment-routing prefix-sid-map advertise-local
fast-reroute per-prefix level-2 all
fast-reroute ti-lfa level-2
microloop avoidance rib-update-delay 10000
!
```



## CHAPTER 7

# SR-TE Per-Flow (Class) ODN and Automated Steering (PCE Delegated)

Table 2: Feature History

Feature Name	Release	Description
SR-TE Per-Flow (Class) ODN and Automated Steering (PCE Delegated)	Cisco IOS XE Amsterdam 17.3.1	This feature lets you steer traffic with SR-TE PFP based on the QoS markings on the packets. The traffic is then switched onto the appropriate path based on the forward classes of the packet.

The SR-TE Per-flow policy (PFP) ODN with auto steering (Per flow ODN/AS) is a mechanism that allows the steering of traffic on an SR policy based on the attributes of the packets. SR-TE Per-flow policy (PFP) ODN with auto steering (Per flow ODN/AS) is a mechanism that allows the steering of traffic on an SR policy based on the attributes of the packet. Packets are classified using Cisco's Modular QoS CLI (MQC) framework and then marked using internal tags known as forward classes (FCs). A Per-Flow Policy (PFP) is then used to route the marked packets based on the mappings between an FC and its corresponding path. This means that the traffic is steered based on its QoS markings and switched onto the appropriate path based on the FC of the packet.

A PFP is identified by <color, endpoint>. It is configured with a per-flow forwarding class table with up to eight entries, with each entry indexed by an FC and points to a Per Destination Policy (PDP) or native RIB path.

- [Restrictions for SR-TE Per-Flow \(Class\) ODN and Automated Steering \(PCE Delegated\)](#), on page 79
- [Configuring SR ODN Per-Flow Policy \(PFP\) AUTO STEERING \(PCE DELEGATED\)](#), on page 81

## Restrictions for SR-TE Per-Flow (Class) ODN and Automated Steering (PCE Delegated)

- 250 PFP+PDP (Combination) is supported.
- Dynamic change in the Quality of Service policy is not supported.
- PIC core over SR-TE tunnel PIC edge is not supported.

- VPLS over SR-TE is not supported.
- IPV6 is not supported.
- 10k VPNV4 prefix limit is supported.
- Effective Cisco IOS XE Bengaluru 17.5.1, ECMP over SR-TE is supported on RSP3.
- Total number of labels supported is 5 (3+2).
- Configure the set forward class to 0 to take default path for non-forward class.
- L3VPN Inter AS Option B for SR PFP is supported.
- Bgp vpnv4 label allocation mode should be on per VRF mode.
- BGP Labeled Unicast (BGP-LU) (RFC 3107) is not supported for SR ODN PFP Auto Steering.
- L2VPN over PFP tunnels is not supported.
- IPv6 services over PFP tunnel is not supported.
- Performance-Measurement over PFP is not supported.
- MPLS Ping or trace-route over PFP is not supported.
- Auto-route announce over PFP or PDP is not supported.
- PIC is not supported over PFP.
- IPv4 over MPLS and MPLS over VPN services are not supported with three and four transport labels.

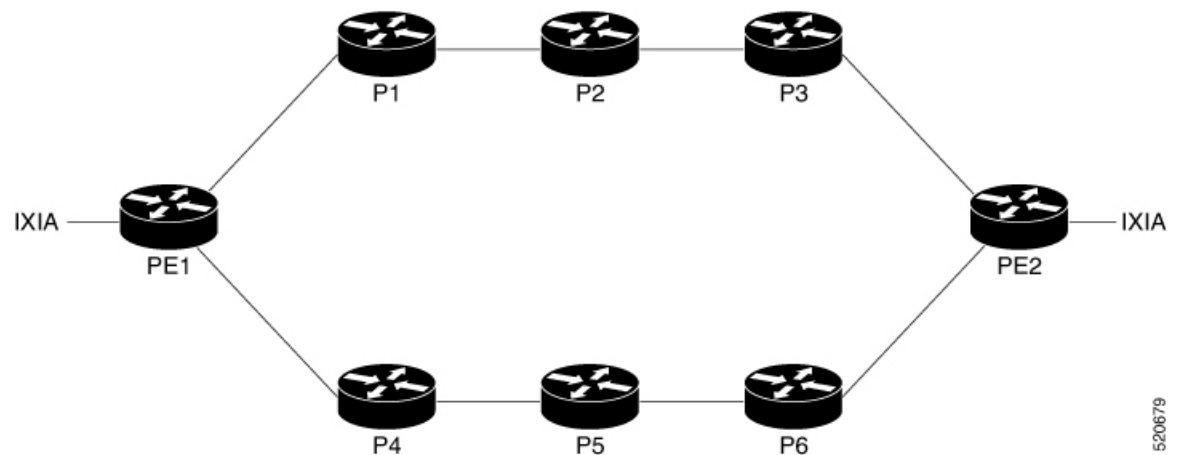
For RSP3:

- HW load balancing does not work with SR PFP template. HW load balancing works based on pre-selection only.
- SR PM will turn into SW timestamping mode if SRTE PFP feature is enabled through template.
- The following Quality of Service must be disabled in this option
  - Short Pipe
  - Vlan Cos Marking
- Timestamping will work at PD SW.
- A download error message is displayed for all unsupported OCE chains.
- PIC core over SR-TE tunnel PIC edge is not supported.
- VPLS over SR-TE is not supported.
- Effective Cisco IOS XE Bengaluru 17.5.1, ECMP over SR-TE is supported on RSP3.
- IPv4 over MPLS and MPLS over VPN services are not supported with three and four transport labels.
- Segment routing in Active or Active Quality of Service template is not supported.
- Short-pipe mode not supported.
- IPV6 is not supported.

- Policy statistics are not supported for 3/4 label SR policy.
- 250 PFP policies and 250 PDP policies are supported.
- The SR PFP and SR five-label push template are mandatory for PFP to work. When using the SR five-label push template, the following features are not supported:
  - OAM down MEP
  - Ethernet loopback
  - RSPAN
- The enable\_egr\_l3vpn\_cm – L3VPN conditional marking is an optional template for PFP.
- Any other template cannot be enabled along with the OAM down MEP, Ethernet loopback, or RSPAN templates.
- LB paths under PDP or PFP tunnels are not supported.

## Configuring SR ODN Per-Flow Policy (PFP) AUTO STEERING (PCE DELEGATED)

Consider the following topology:



Perform the following steps to configure ODN for PFP:

1. Configure Quality of Service on PE1.

```

class-map DSCP
  match DSCP AF41
  
```

- Set forward class on the class map.

```

policy-map per-flow
  class DSCP
    set forward-class 1
  
```

- Attach the policy map on the corresponding interface.

```
interface TenGigabitEthernet0/1/0
service-policy input per-flow
```

## 2. Configure SR-TE PFP on PE1.

- Set forward class on PFP.

```
policy PER_FLOW
color 4500 end-point 5.5.5.5
candidate-paths
  preference 2
  per-flow
    forward-class 0 color 100
    forward-class 1 color 330
    forward-class 2 color 580
    forward-class 3 color 3010
    forward-class 4 color 2305
    forward-class 5 color 8363
    forward-class 6 color 9000
    forward-class 7 color 4294967295
```

- Attach the segment list to PDP.

```
policy perflow_pdp
color 100 end-point 5.5.5.5
candidate-paths
  preference 2
  explicit segment-list srtel weight 10
  !
constraints
  segments
    dataplane mpls
```

- Set segment list to SR-TE.

```
segment-routing traffic-eng
segment-list name srtel
  index 1 mpls label 16002
  index 2 mpls label 16005
```

## 3. Configure SR-TE PFP on PE2.

```
ip prefix-list pfp seq 5 permit 35.0.0.0/16 le 32
```

- Attach route-map to PFP.

```
route-map pfp permit 10
match ip address prefix-list pfp
set extcommunity color 4500
```

- Activate BGP routes.

```
router bgp 100
!
address-family vpnv4
neighbor 1.1.1.1 activate
neighbor 1.1.1.1 send-community extended
neighbor 1.1.1.1 route-map pfp out
```





## CHAPTER 8

# SR-TE On Demand LSP

Effective Cisco IOS XE Everest 16.6.1, SR TE On demand LSP feature provides the ability to connect Metro access rings via a static route to the destination. The static route is mapped to an explicit path that triggers an on demand LSP to the destination. The SR TE On demand LSP feature will be used to transport the VPN services between the Metro access rings.



### Note

SR-TE On Demand LSP is available only on the Cisco RSP2 Module.

- [Restrictions for SR-TE On Demand LSP, on page 83](#)
- [Information About SR-TE On Demand LSP, on page 83](#)
- [How to Configure SR-TE On Demand LSP, on page 84](#)

## Restrictions for SR-TE On Demand LSP

- Segment-Routing auto tunnel static route does not support ECMP.
- Metrics for IP explicit path and administrative distance change for auto tunnel SRTE static route is not supported.
- MPLS Traffic Engineering (TE) Nonstop Routing (NSR) must be configured on the active route processor (RP) for Stateful Switchover (SSO). This is because SR static auto tunnel fails to come up after SSO, unless the static route auto tunnel configuration is removed and reconfigured.
- IP unnumbered interfaces do not support dynamic path.
- When using IP unnumbered interfaces, you cannot specify next hop address as an explicit path index. It should be a node address or a label.

## Information About SR-TE On Demand LSP

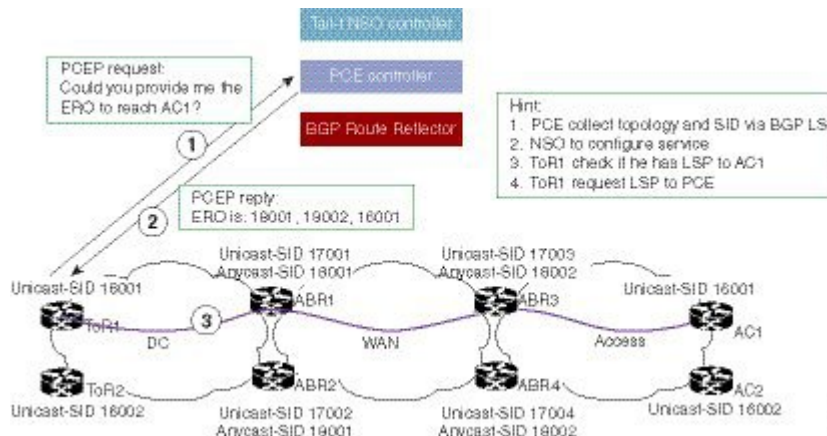
The SR TE On demand LSP feature provides the ability to connect Metro access rings via a static route to the destination.

### SR-TE: Setup LSP as Static Route

Agile Carrier Ethernet (ACE) solution leverages Segment Routing-based transport for consolidated VPN services. In metro rings architecture, the access rings do not share their routing topologies with each other.

The SR TE On demand LSP feature provides the ability to connect Metro access rings via a static route to the destination. The static route is mapped to an explicit path and that will trigger an on demand LSP to the destination. The SR TE On demand LSP feature is used to transport the VPN services between the Metro access rings.

Figure 7: Inter-Metro LSP in ACE Solution



Inter-Metro LSPs have the following aspects:

- The source packet may not know the IP address of the destination device.
- Existing segment routing features are applicable for LSPs.
- The binding SID helps in steering the traffic in the SR-TE tunnel. In other words, ingress MPLS packet with the binding SID will be forwarded through the specific SR-TE tunnel.

## Static SRTE over Unnumbered Interfaces

As explained in the previous section, you can set up LSP as static route to create an auto tunnel by specifying an IP explicit path.

The explicit path is a combination of IP addresses (or) IP address and labels. You can also configure the static SRTE tunnel over unnumbered interfaces. There are few restrictions for unnumbered interfaces against numbered interfaces.

- You must specify the node IP address, not the next hop interface address in the ip-explicit path option.
- You must not specify adjacency SID in the explicit path option. In short, the explicit path option should contain only the node IP address (/32 mask) and prefix SID labels.

## How to Configure SR-TE On Demand LSP

Perform the following steps to configure SR-TE On Demand LSP.

### Configuring LSP as Static Route

To avoid packet drop after RP switchover with SR TE, it is recommended to use the following command:

```
mpls traffic-eng nsr
```

If ISIS is configured, use the following command:

```
router isis
 nsf cisco
 nsf interval 0
```

## Enabling Segment Routing Auto Tunnel Static Route

Perform this task to configure auto tunnel static route as follows:

- Configure IP explicit path
- Associate the auto tunnel with an IP explicit path with a static route
- Enable peer-to-peer (P2P) auto tunnel service

```
ip explicit-path name path1
 index 1 next-label 16002
 index 2 next-label 16006
 exit
mpls traffic-eng auto-tunnel p2p
mpls traffic-eng auto-tunnel p2p config unnumbered-interface loopback0
mpls traffic-eng auto-tunnel p2p tunnel-num min 10 max 100
ip route 172.16.0.1 255.240.0.0 segment-routing mpls path name path1
```

## Verifying Segment Routing Auto-Tunnel Static Route

The command **show mpls traffic-eng service summary** displays all registered TE service clients and statistics that use TE auto tunnel.

```
Device# show mpls traffic-eng service summary
Service Clients Summary:
  Client: BGP TE
    Client ID           :0
    Total P2P tunnels   :1
    P2P add requests    :6
    P2P delete requests :5
  P2P add falis        :0
    P2P delete falis    :0
    P2P notify falis    :0
    P2P notify succs    :12
    P2P replays         :0
  Client: ipv4static
    Client ID           :1
    Total P2P tunnels   :1
    P2P add requests    :6
    P2P delete requests :5
    P2P add falis       :0
    P2P delete falis    :0
    P2P notify falis    :0
    P2P notify succs    :85
    P2P replays         :0
```

The command **show mpls traffic-eng auto-tunnel p2p** displays the peer-to-peer (P2P) auto tunnel configuration and operation status.

The command **show mpls traffic-eng tunnel summary** displays the status of P2P auto tunnel.

```
Device# show mpls traffic-eng auto-tunnel p2p
State: Enabled
  p2p auto-tunnels: 2 (up: 2, down: 0)
  Default Tunnel ID Range: 62336 - 64335
  Config:
unnumbered-interface: Loopback0 Tunnel ID range: 1000 - 2000
```

The command **show mpls traffic-eng tunnel summary** displays the status of P2P auto tunnel.

```
Device# show mpls traffic-eng tunnel summary
Signalling Summary:
  LSP Tunnels Process:
  Passive LSP Listener:
  RSVP Process:
  Forwarding:
  auto-tunnel:
running
running
running
enabled
p2p Enabled (1), id-range:1000-2000
Periodic reoptimization:
Periodic FRR Promotion:
Periodic auto-bw collection:
SR tunnel max label push:
P2P:
every 3600 seconds, next in 1265 seconds
Not Running
every 300 seconds, next in 66 seconds
13 labels
  Head: 11 interfaces, 5234 active signalling attempts, 1 established
        5440 activations, 206 deactivations
        1821 failed activations
        0 SSO recovery attempts, 0 SSO recovered
  Midpoints: 0, Tails: 0
P2MP:
  Head: 0 interfaces, 0 active signalling attempts, 0 established
        0 sub-LSP activations, 0 sub-LSP deactivations
        0 LSP successful activations, 0 LSP deactivations
        0 SSO recovery attempts, LSP recovered: 0 full, 0 partial, 0 fail
  Midpoints: 0, Tails: 0
Bidirectional Tunnel Summary:
  Tunnel Head: 0 total, 0 connected, 0 associated, 0 co-routed
LSPs Head:
LSPs Mid:
LSPs Tail:
0 established, 0 proceeding, 0 associated, 0 standby
0 established, 0 proceeding, 0 associated, 0 standby
0 established, 0 proceeding, 0 associated, 0 standby
AutoTunnel P2P Summary:
  ipv4static:
    Tunnels: 1 created, 1 up, 0 down
  Total:
    Tunnels: 1 created, 1 up, 0 down
```

The command **show ip explicit-path** displays the configured IP explicit paths.

```
Device# show ip explicit-path
```

```

PATH path1 (strict source route, path complete, generation 5)
  1: next-label 16055
  2: next-label 16070
  3: next-label 16010
PATH path2 (strict source route, path complete, generation 9)
  1: next-label 16070
  2: next-label 16010
Device# show running-config | i ip route
ip route 10.10.10.10 255.255.255.255 segment-routing mpls path name path1 verbatim
ip route 11.11.11.1 255.255.255.255 segment-routing mpls path name path2 verbatim
ip route 11.11.11.2 255.255.255.255 segment-routing mpls path name path2 verbatim

```

The command **show mpls traffic-eng tunnel auto-tunnel detail** only displays TE service auto tunnel.

```

Device# show mpls traffic-eng tunnel auto-tunnel detail
P2P TUNNELS/LSPs:

Name: R1-RSP2_t4000                                (Tunnel4000) Destination: 0.0.0.0 Ifhandle: 0x23
(auto-tunnel for ipv4static)
Status:
  Admin: up      Oper: up      Path: valid      Signalling: connected
  path option 1, (SEGMENT-ROUTING) type explicit (verbatim) path1 (Basis for Setup)

Config Parameters:
  Bandwidth: 0      kbps (Global) Priority: 7 7 Affinity: 0x0/0xFFFF
  Metric Type: TE (default)
  Path Selection:
    Protection: any (default)
  Path-selection Tiebreaker:
    Global: not set Tunnel Specific: not set Effective: min-fill (default)
  Hop Limit: disabled [ignore: Verbatim Path Option]
  Cost Limit: disabled
  Path-invalidation timeout: 10000 msec (default), Action: Tear
  AutoRoute: disabled LockDown: disabled Loadshare: 0 [0] bw-based
  auto-bw: disabled
  Fault-OAM: disabled, Wrap-Protection: disabled, Wrap-Capable: No
Active Path Option Parameters:
  State: explicit path option 1 is active
  BandwidthOverride: disabled LockDown: disabled Verbatim: enabled

History:
  Tunnel:
    Time since created: 7 days, 22 hours, 48 minutes
    Time since path change: 7 days, 22 hours, 47 minutes
    Number of LSP IDs (Tun_Instances) used: 9
    Current LSP: [ID: 9]
    Uptime: 7 days, 22 hours, 47 minutes
  Tun_Instance: 9
  Segment-Routing Path Info (IGP information is not used)
    Segment0[First Hop]: 0.0.0.0, Label: 16055
    Segment1[ - ]: Label: 16070
    Segment2[ - ]: Label: 16010

```

The command **show mpls traffic-eng tunnel brief** displays auto tunnel information.

```

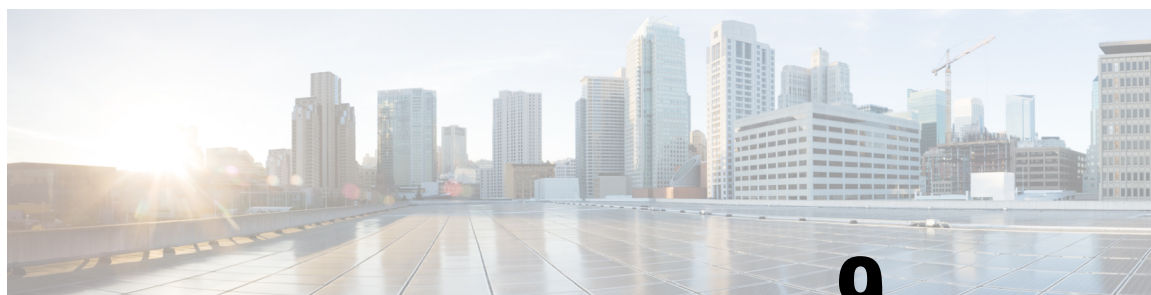
Device# show mpls traffic-eng tunnel brief
Signalling Summary:
  LSP Tunnels Process: running
  Passive LSP Listener: running
  RSVP Process: running
  Forwarding: enabled

```

```

auto-tunnel:
p2p    Enabled (2), id-range:1000-2000
Periodic reoptimization: every 3600 seconds, next in 406 seconds
Periodic FRR Promotion: Not Running
Periodic auto-bw collection: every 300 seconds, next in 107 seconds
SR tunnel max label push: 13 labels
P2P TUNNELS/LSPs:
TUNNEL NAME    DESTINATION    UP IF    DOWN IF    STATE/PROT
R1_t1          66.66.66.66    -        -          up/down
R1_t2          66.66.66.66    -        -          up/down
R1_t3          66.66.66.66    -        -          up/down
R1_t10         66.66.66.66    -        -          up/down
SBFD tunnel    33.33.33.33    -        -          up/down
SBFD Session configured: 1 SBFD sessions UP: 1

```



## CHAPTER 9

# SR-TE On-Demand Next Hop

When redistributing routing information across domains, provisioning of multi-domain services (Layer 2 VPN and Layer 3 VPN) has its own complexity and scalability issues. Effective Cisco IOS XE Everest 16.6.1, On-Demand Next Hop (ODN) triggers delegation of computation of an end-to-end LSP to a PCE controller including constraints and policies without doing any redistribution. It then installs the reapplied multi-domain LSP for the duration of the service into the local forwarding information base (FIB).



**Note** SR-TE On-Demand Next Hop is available only on the Cisco RSP2 Module.

- [Restrictions for SR-TE On-Demand Next Hop, on page 89](#)
- [Information About SR-TE On-Demand Next Hop, on page 89](#)
- [SR-TE Policy, Color Extended Community, Affinity Constraint, and Disjointness Constraint, on page 93](#)
- [How to Configure Segment Routing On-Demand Next Hops—Layer 3 and Layer 3 VPN, on page 95](#)
- [Verifying Segment Routing On-Demand Next Hop—Layer 3 and Layer 3 VPN, on page 99](#)
- [Configuring Color Extended Community, Affinity Constraint, and Disjointness Constraint, on page 103](#)
- [Verifying SR-TE ODN Color Extended Community, Affinity Constraint, and Disjointness Constraint, on page 106](#)
- [Troubleshooting the SR-TE ODN Color Extended Community, Affinity Constraint, and Disjointness Constraint, on page 110](#)

## Restrictions for SR-TE On-Demand Next Hop

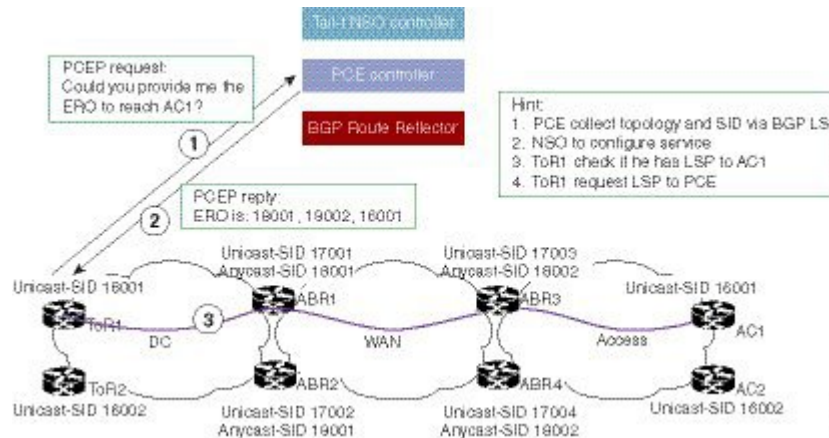
- ODN anycast SID is not supported.
- ODN for IPv6 is not supported.
- SR ODN Tunnel is *not* supported on BGP NSR. It is only supported on BGP NSF.

## Information About SR-TE On-Demand Next Hop

On-Demand Next hop leverages upon BGP Dynamic SR-TE capabilities and adds the path computation (PCE) ability to find and download the end to end path based on the requirements. ODN triggers an SR-TE auto-tunnel

based on the defined BGP policy. As shown in the following figure, an end-to-end path between ToR1 and AC1 can be established from both ends based on low latency or other criteria for VRF (L3VPN) or IPv4 services. The work-flow for ODN is summarized as follows:

**Figure 8: ODN Operation**



1. PCE controller collects topology and SIDs information via BGP Link State (BGP-LS). For more information on BGP-LS, refer [BGP Link-State](#).
2. If NSO controller is enabled, it configures L3VPN VRF or IPv4 prefixes and requests are sent to ToR1 and AC1.
3. ToR1 and AC1 checks if an LSP toward each other exists. If not, a request is sent to the PCE controller to compute that SR-TE path that matches SR-TE policy that is carried through BGP.
4. PCE controller computes the path and replies with a label stack (18001, 18002, 16001, example in ToR1).
5. ToR1 and AC1 create an SR-TE auto-tunnel and reply to the NSO controller indicating that the LSP for VRF or IPv4 is up and operational.

## Fast Convergence Default Optimize

The fast convergence "default optimize" feature modifies the default settings of all the protocols to recommended defaults for fast convergence. To revert the defaults to pre-fast-convergence settings for both IS-IS and OSPF, **routing-default-optimize** command is used. This command sends signals to IS-IS and OSPF and modifies the default configuration for these protocols.

By default, the fast convergence setting is enabled which means when you upgrade the software, you can automatically see the new behavior. This makes easier integration of the devices in a multivendor deployment and reduces support cases for poor convergence.

When default optimize is disabled, existing protocol default configuration is used. When default optimize is enabled, new protocol defaults are used. The show running configurations does not display configuration lines for default settings even when default settings are being used.

A configuration of a protocol overrides the default, but a change to default optimize does not override any configuration.

The following is the sample output of **spf-interval** command in IS-IS:



```
Device(config-if)# router isis
Device(config-router)# spf-interval 10 5500 5500
```

If a non-default value is configured, it is displayed in show running configuration output:

```
Device(config-router)# spf-interval 5 50 200
Device(config-router)# do show run | inc spf-interval
spf-interval 5 50 200
```

You can revert to the default values by configuring the default values or by removing the non-default configuration.

## Default Optimize Values for IS-IS

The following table summarizes the configuration impacted by default optimize:

IS-IS command	Parameters	Default optimize disabled	Default optimize enabled
fast-flood			
	number of LSPs flooded back-to-back	Disabled	10
spf-interval			
	Initial (milliseconds)	5500	50
	Secondary (milliseconds)	5500	200
	Max (seconds)	10	5
prc-interval			
	Initial (milliseconds)	2000	50
	Secondary (milliseconds)	5000	200
	Max (seconds)	5	5
lsp-gen-interval			
	Initial (milliseconds)	50	50
	Secondary (milliseconds)	5000	200
	Max (seconds)	5	5
log-adjacency-changes		disabled	enabled

## Default Optimize Values for OSPF

The following table summarizes the configuration impacted by default optimize for OSPFv2/v3:

OSPF command	Parameters	Default optimize disabled	Default optimize enabled
timers throttle spf			
	Initial (milliseconds)	5500	50
	Secondary (milliseconds)	10000	200
	Max (seconds)	10	5
timers throttle lsa all			
	Initial (milliseconds)	0	50
	Secondary (milliseconds)	5000	200
	Max (seconds)	5	5
timers lsa arrival			
	milliseconds	1000	100

The following is the sample output of show ip ospf command for OSPFv2 with the default-optimize values.

```

Device# show ip ospf
Routing Process "ospf 10" with ID 1.1.1.1
Start time: 00:00:01.471, Time elapsed: 03:00:34.706
Supports only single TOS(TOS0) routes
Supports opaque LSA
Supports Link-local Signaling (LLS)
Supports area transit capability
Supports NSSA (compatible with RFC 3101)
Supports Database Exchange Summary List Optimization (RFC 5243)
Event-log enabled, Maximum number of events: 1000, Mode: cyclic
Router is not originating router-LSAs with maximum metric
Initial SPF schedule delay 50 msecs
Minimum hold time between two consecutive SPF's 200 msecs
Maximum wait time between two consecutive SPF's 5000 msecs
Incremental-SPF disabled
Initial LSA throttle delay 50 msecs
Minimum hold time for LSA throttle 200 msecs
Maximum wait time for LSA throttle 5000 msecs
Minimum LSA arrival 100 msecs
LSA group pacing timer 240 secs
Interface flood pacing timer 33 msecs
Retransmission pacing timer 66 msecs
EXCHANGE/LOADING adjacency limit: initial 300, process maximum 300
Number of external LSA 18. Checksum Sum 0x075EB2
Number of opaque AS LSA 0. Checksum Sum 0x000000
Number of DCbitless external and opaque AS LSA 0
Number of DoNotAge external and opaque AS LSA 0
Number of areas in this router is 1. 1 normal 0 stub 0 nssa
Number of areas transit capable is 0
External flood list length 0
IETF NSF helper support enabled
Cisco NSF helper support enabled
Reference bandwidth unit is 100 mbps
  Area BACKBONE(0)
    Number of interfaces in this area is 4 (2 loopback)
    Area has RRR enabled

```

```

Area has no authentication
SPF algorithm last executed 02:27:23.736 ago
SPF algorithm executed 20 times
Area ranges are
Number of LSA 94. Checksum Sum 0x321DCF
Number of opaque link LSA 0. Checksum Sum 0x000000
Number of DCbitless LSA 0
Number of indication LSA 0
Number of DoNotAge LSA 0
Flood list length 0

```

The following is the sample output of **show ospf** command for OSPFv3 with the default-optimize values.

```

Device# show ospfv3
OSPFv3 10 address-family ipv6
Router ID 11.11.11.11
Supports NSSA (compatible with RFC 3101)
Supports Database Exchange Summary List Optimization (RFC 5243)
Event-log enabled, Maximum number of events: 1000, Mode: cyclic
Router is not originating router-LSAs with maximum metric
Initial SPF schedule delay 50 msec
Minimum hold time between two consecutive SPF's 200 msec
Maximum wait time between two consecutive SPF's 5000 msec
Initial LSA throttle delay 50 msec
Minimum hold time for LSA throttle 200 msec
Maximum wait time for LSA throttle 5000 msec
Minimum LSA arrival 100 msec
LSA group pacing timer 240 secs
Interface flood pacing timer 33 msec
Retransmission pacing timer 66 msec
Retransmission limit dc 24 non-dc 24
EXCHANGE/LOADING adjacency limit: initial 300, process maximum 300
Number of external LSA 0. Checksum Sum 0x000000
Number of areas in this router is 1. 1 normal 0 stub 0 nssa
Graceful restart helper support enabled
Reference bandwidth unit is 100 mbps
RFC1583 compatibility enabled
Area BACKBONE(0)
  Number of interfaces in this area is 2
  SPF algorithm executed 7 times
  Number of LSA 3. Checksum Sum 0x012426
  Number of DCbitless LSA 0
  Number of indication LSA 0
  Number of DoNotAge LSA 0
  Flood list length 0

```

## SR-TE Policy, Color Extended Community, Affinity Constraint, and Disjointness Constraint

Effective Cisco IOS XE Gibraltar 16.11.1, the router supports:

- MPLS TE new SR-TE policy command—**segment-routing traffic-eng**
- Color-extended community
- Affinity constraints
- Disjointness constraints

## SR-TE Policy Command

Effective Cisco IOS XE Gibraltar 16.11.1, a new command (**segment-routing traffic-eng**) is added to configure the SR policy under segment routing. For information on using this command, see the “*Configuring Color-Extended Community*” section.

## Color Extended Community

In earlier releases, the router created segment routed Traffic Engineering (SR-TE) tunnels based on a tunnel-profile or attribute set. As part of this functionality, an inbound route-map with a “match community” and “set attribute-set” was added on the ingress node and the route-map matched against communities received in the BGP updates. A BGP update with a matching community would initiate an SR-TE tunnel for the nexthop TE-profile.

However, effective Cisco IOS XE Gibraltar 16.11.1, ‘color extended’ community is now added as follows:

- An SR-TE policy is created on the ingress router for the Color-Endpoint pair.
- The egress router adds the ‘color extended’ community to the BGP updates that require a Traffic-Engineered path.

## Affinity Constraint

Affinity is a 32-bit constraint used by the PCE and PCALC for calculating paths that take the “affinity constraint” into account.

Affinity constraints let you assign, or map, color names for path affinities. After mappings are defined, the attributes can be referred to by the corresponding color name in the command.

Affinity maps are used to map operator-defined color names to a bit position in the affinity bitmap

Supported Affinity constraints are:

- include-all—indicates that constrained shortest path first (CSPF) includes a link when calculating a path, only if each link administrative group bit has the same name as each affinity bit.
- include-any—indicates that CSPF includes a link when calculating a path, if at least one link administrative group bit has the same name as an affinity bit.
- exclude-any—indicates that CSPF excludes a link when calculating a path, if any link administrative group bit has the same name as an affinity bit.

## Disjointness Constraint

Disjointness is used to describe two or more services that must be completely disjoint of each other. Disjointness is useful for providing traffic flow redundancy in the network.

Disjointness is controlled by the PCE. The PCE learns of the network topology through an IGP (OSPF or IS-IS) through the BGP-LS protocol and is capable of computing paths based on the IGP or TE metric.

The PCE uses the disjoint policy to compute two lists of segments that steer traffic from the source node to the destination node along disjoint paths. Disjoint paths can originate from either the same or different head-ends.

A "disjoint level" refers to the type of resources that should not be shared by the two computed paths. The PCE supports the following disjoint path computations:

- Link
- Node
- Shared risk link group (SRLG)

When the first request is received from Path Computation client (PCC) or an ingress node, with a given disjoint-group ID, a list of segments is computed based on the metric requested, encoding the shortest path from source to destination.

When the second request is received with the same disjoint-group ID, based on the information received in both requests, the PCE computes two disjoint paths from the source to the destination.

Both paths are computed at the same time. The shortest list of segments is calculated to steer traffic on the computed paths.

## Limitations

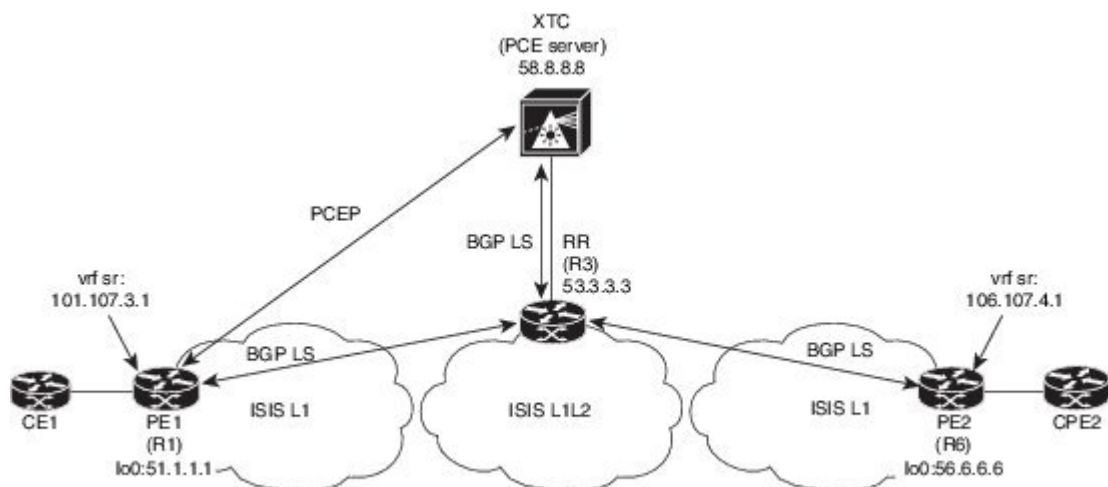
- The maximum number of SR-TE ODN policies supported is 500.
- Static SR-TE policy is not supported.
- SR-TE policy goes down only after 60 seconds when the connection to PCEP is disconnected due to any reason.
- 6PE and 6VPE services are supported only if the number of transport labels used for SR-TE policy does not exceed two labels.

# How to Configure Segment Routing On-Demand Next Hops—Layer 3 and Layer 3 VPN

## Configuring Segment Routing On-Demand Next Hops—Layer 3 and Layer 3 VPN

Perform the following steps to configure on-demand next hop for SR-TE. The following figure is used as a reference to explain the configuration steps.

Figure 9: ODN Auto-Tunnel Setup



36/02/23

1. Configure the router (R6 tailend) with VRF interface.

```
vrf definition sr
rd 100:100
route-target export 100:100
route-target import 100:100
!
address-family ipv4
exit-address-family
!
address-family ipv6
exit-address-family
!
interface gigabitEthernet0/0/11
vrf forwarding sr
ip address 30.30.30.1 255.255.255.0
```

2. Tags VRF prefix with BGP community on R6 (tailend).

```
route-map BGP_TE_MAP permit 9
match ip address L3VPN_ODN_ROUTES
set community 3276850
ip access-list extended L3VPN_ODN_ROUTES
permit ip 30.30.30.1 255.255.0.0 any
```

3. Enable BGP on R6 (tailend) and R1 (headend) to advertise and receive VRF SR prefix and match on community set on R6 (tailend).

```
router bgp 100 <-- This BGP configuration applies to the tailend
bgp router-id 172.16.0.1
bgp log-neighbor-changes
bgp graceful-restart
no bgp default ipv4-unicast
neighbor 10.0.0.2 remote-as 100
neighbor 10.0.0.2 update-source Loopback0
address-family ipv4
neighbor 10.0.0.2 activate
neighbor 10.0.0.2 send-community both
```

```

    neighbor 10.0.0.2 next-hop-self
  exit-address-family
  address-family vpnv4
    neighbor 10.0.0.2 activate
    neighbor 10.0.0.2 send-community both
    neighbor 10.0.0.2 route-map BGP_TE_MAP out
  exit-address-family
  address-family link-state link-state
    neighbor 10.0.0.2 activate
  exit-address-family
  address-family ipv4 vrf sr
    redistribute connected
  exit-address-family
  route-map BGP_TE_MAP permit 9
    match ip address traffic
    set community 3276850
  ip access-list extended traffic
    permit ip 10.0.0.1 255.255.0.0 any
  router bgp 100 <-- This BGP configuration applies to the headend
    bgp router-id 192.168.0.2
    bgp log-neighbor-changes
    bgp graceful-restart
    no bgp default ipv4-unicast
    neighbor 10.0.0.2 remote-as 100
    neighbor 10.0.0.2 update-source Loopback0
  address-family ipv4
    neighbor 10.0.0.2 activate
    neighbor 10.0.0.2 send-community both
    neighbor 10.0.0.2 next-hop-self
  exit-address-family
  address-family vpnv4
    neighbor 10.0.0.2 activate
    neighbor 10.0.0.2 send-community both
    neighbor 10.0.0.2 route-map BGP_TE_MAP in
  exit-address-family
  address-family link-state link-state
    neighbor 10.0.0.2 activate
  exit-address-family
  address-family ipv4 vrf sr
    redistribute connected
  exit-address-family
  route-map BGP_TE_MAP permit 9
    match community 1
    set attribute-set BGP_TE5555
  ip community-list 1 permit 3276850
  mpls traffic-eng lsp attributes BGP_TE5555
    path-selection metric igp
  pce

```

#### 4. Enable route-map or attribute set on headend (R1).

```

route-map BGP_TE_MAP permit 9
match community 1
set attribute-set BGP_TE5555
ip community-list 1 permit 3276850
mpls traffic-eng lsp attributes BGP_TE5555
path-selection metric igp
pce
end

```

#### 5. Enable PCE and auto-tunnel configurations on R1.

```

mpls traffic-eng tunnels

```

```
mpls traffic-eng pcc peer 10.0.0.3 source 10.0.0.4 precedence 255
mpls traffic-eng auto-tunnel p2p tunnel-num min 2000 max 5000
```

6. Enable all core links with SR-TE configurations and ensure that they are enabled as point-to-point interfaces.

```
mpls traffic-eng tunnels
interface GigabitEthernet0/2/0
ip address 101.102.6.1 255.255.255.0
ip router isis 1
mpls traffic-eng tunnels
isis network point-to-point
interface GigabitEthernet0/3/1
vrf forwarding sr
ip address 101.107.3.1 255.255.255.0
negotiation auto
end
```

7. Enable R3 (RR) to advertise TED to the PCE server via BGP-LS.

```
router isis 1
net 49.0002.0000.0000.0003.00
ispsf level-1-2
metric-style wide
nsf cisco
nsf interval 0
distribute link-state
segment-routing mpls
segment-routing prefix-sid-map advertise-local
redistribute static ip level-1-2
mpls traffic-eng router-id Loopback0
mpls traffic-eng level-1
mpls traffic-eng level-2
router bgp 100
bgp router-id 10.0.0.2
bgp log-neighbor-changes
bgp graceful-restart
no bgp default ipv4-unicast
neighbor 10.0.0.3 remote-as 100
neighbor 10.0.0.3 update-source Loopback0
address-family ipv4
neighbor 10.0.0.3 activate
exit-address-family
```

8. Enable PCE server configuration and verify that BGP-LS session is properly established with RR.

```
Device# show bgp li li summary
BGP router identifier 10.0.0.3, local AS number 100
BGP generic scan interval 60 secs
Non-stop routing is enabled
BGP table state: Active
Table ID: 0x0 RD version: 1436
BGP main routing table version 1436
BGP NSR Initial initsync version 1 (Reached)
BGP NSR/ISSU Sync-Group versions 0/0
BGP scan interval 60 secs
BGP is operating in STANDALONE mode.
Process RcvTblVer bRIB/RIB LabelVer ImportVer SendTblVer StandbyVer
Speaker 1436 1436 1436 1436 1436
0
Neighbor Spk AS MsgRcvd MsgSent TblVer InQ OutQ Up/Down St/PfxRcd
10.0.0.2 0 100 19923 17437 1436 0 0 1w2d 103
Device# show pce ipv4 topo | b Node 3
```



```

Node 3
  TE router ID: 10.0.0.2
  Host name: R3
  ISIS system ID: 0000.0000.0003 level-1
  ISIS system ID: 0000.0000.0003 level-2
  Prefix SID:
    Prefix 10.0.0.2, label 20011 (regular)

```



**Note** The above **show** command is run only on the PCE server, for example the Cisco ASR 9000 Series Aggregation Services Routers.

## Verifying Segment Routing On-Demand Next Hop—Layer 3 and Layer 3 VPN

The ODN verifications are based on L3VPN VRF prefixes.

1. Verify that PCEP session between R1 (headend and PCE server) is established.

```

Device# show pce client peer
PCC's peer database:
-----
Peer address: 10.0.0.3 (best PCE)
State up
Capabilities: Stateful, Update, Segment-Routing

```

2. Verify that PCEP session is established between all the peers (PCCs).

```

Device# show pce ipv4 peer
PCE's peer database:
-----
Peer address: 10.0.0.4
State: Up
Capabilities: Stateful, Segment-Routing, Update
Peer address: 172.16.0.5
State: Up
Capabilities: Stateful, Segment-Routing, Update

```



**Note** The above **show** command is run only on the PCE server, for example the Cisco ASR 9000 Series Aggregation Services Routers.

3. Verify that R1 (headend) has no visibility to R6 loopback address.

```

Device# show ip route 192.168.0.1
% Network not in table

```

4. Verify that VRF prefix is injected via MP-BGP in R1 VRF SR routing table.

```

Device# show ip route vrf sr
Routing Table: sr

```

```

Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
      N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
      E1 - OSPF external type 1, E2 - OSPF external type 2
      i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
      ia - IS-IS inter area, * - candidate default, U - per-user static route
      o - ODR, P - periodic downloaded static route, H - NHRP, l - LISP
      a - application route
      + - replicated route, % - next hop override, p - overrides from PfR
Gateway of last resort is not set
C L
B
10.0.0.6/8 is variably subnetted, 2 subnets, 2 masks
    10.0.0.7/24 is directly connected, GigabitEthernet0/3/1
    10.0.0.7/32 is directly connected, GigabitEthernet0/3/1
    10.0.0.8/24 is subnetted, 1 subnets
    10.0.0.9 [200/0] via binding label: 865, 4d21h

```

5. Verify that BGP is associating properly the policy and binding SID with the VRF prefix.

```

Device# show ip bgp vpnv4 vrf sr 106.107.4.0
BGP routing table entry for 100:100:106.107.4.0/24, version 3011
Paths: (1 available, best #1, table sr)
  Not advertised to any peer
  Refresh Epoch 4
  Local
    192.168.0.1 (metric 10) (via default) from 10.0.0.2 (10.0.0.2)
      Origin incomplete, metric 0, localpref 100, valid, internal, best
      Community: 3276850
      Extended Community: RT:100:100
      Originator: 192.168.0.1, Cluster list: 10.0.0.2
      mpls labels in/out nlabel/1085
      binding SID: 865 (BGP_TE5555)
      rx pathid: 0, tx pathid: 0x0

```

6. Verify binding label association with VRF prefix.

```

Device# show ip route vrf sr 106.107.4.0
Routing Table: sr
Routing entry for 106.107.4.0/24
  Known via "bgp 100", distance 200, metric 0, type internal
  Routing Descriptor Blocks:
    * Binding Label: 865, from 10.0.0.2, 4d22h ago
      Route metric is 0, traffic share count is 1
      AS Hops 0
      MPLS label: 1085
      MPLS Flags: NSF

```

7. Verify that VRF prefix is forwarded via ODN auto-tunnel.

```

Device# show ip cef label-table
Label Next Hop Interface
0 no route
865 attached Tunnel2000
Device# show ip cef vrf sr 106.107.4.0 detail
10.0.0.8/24, epoch 15, flags [rib defined all labels]
  recursive via 865 label 1085
  attached to Tunnel2000

```

8. Verify ODN auto-tunnel status.

```

Device# show mpls traffic-eng tunnels

```

```

P2P TUNNELS/LSPs:
Name: R1_t2000 (Tunnel2000) Destination: 192.168.0.1 Ifhandle: 0x6F5
(auto-tunnel for BGP TE)
Status:
Admin: up Oper: up Path: valid Signalling: connected---.. auto-tunnel
2000
path option 1, (SEGMENT-ROUTING) (PCE) type dynamic (Basis for Setup, path weight
10)
Config Parameters:
Bandwidth: 0 kbps (Global) Priority: 7 7 Affinity: 0x0/0xFFFF
Metric Type: IGP (interface)
Path Selection:
Protection: any (default)
Path-selection Tiebreaker:
Global: not set Tunnel Specific: not set Effective: min-fill (default)
Hop Limit: disabled
Cost Limit: disabled
Path-invalidation timeout: 10000 msec (default), Action: Tear
AutoRoute: disabled LockDown: disabled Loadshare: 0 [0] bw-based auto-bw: disabled
Attribute-set: BGP_TE5555--- attribute-set
Fault-OAM: disabled, Wrap-Protection: disabled, Wrap-Capable: No
Active Path Option Parameters:
State: dynamic path option 1 is active
BandwidthOverride: disabled LockDown: disabled Verbatim: disabled
PCEP Info:
Delegation state: Working: yes Protect: no
Working Path Info:
Request status: processed
Created via PCRep message
PCE metric: 30, type: IGP
Reported paths:
Tunnel Name: Tunnel2000_w
LSPs:
LSP[0]:
from PCE server: 10.0.0.3-- via PCE server
source 10.0.0.4, destination 192.168.0.1, tunnel ID 2000, LSP ID 1
State: Admin up, Operation active
Binding SID: 865
Setup type: SR
Bandwidth: requested 0, used 0
LSP object:
PLSP-ID 0x807D0, flags: D:0 S:0 R:0 A:1 O:2
Metric type: IGP, Accumulated Metric 0
ERO:
SID[0]: Adj, Label 2377, NAI: local 101.102.6.1 remote 10.0.0.10
SID[1]: Unspecified, Label 17, NAI: n/a
SID[2]: Unspecified, Label 20, NAI: n/a
History:
Tunnel:
Time since created: 4 days, 22 hours, 21 minutes
Time since path change: 4 days, 22 hours, 21 minutes
Number of LSP IDs (Tun_Instances) used: 1
Current LSP: [ID: 1]
Uptime: 4 days, 22 hours, 21 minutes
Tun_Instance: 1
Segment-Routing Path Info (isis level-1)
Segment0[Link]: 101.102.6.1 - 10.0.0.10, Label: 2377
Segment1[ - ]: Label: 17
Segment2[ - ]: Label: 20

```

## 9. Verify ODN auto-tunnel LSP status on R1 (headend).

```

Device# show pce client lsp brief
PCC's tunnel database:

```

```

-----
Tunnel Name: Tunnel2000_w
LSP ID 1
Tunnel Name: Tunnel2000_p
R1# sh pce client lsp detail
PCC's tunnel database:
-----
Tunnel Name: Tunnel2000_w
LSPs:
LSP[0]:
  source 10.0.0.4, destination 192.168.0.1, tunnel ID 2000, LSP ID 1
  State: Admin up, Operation active
  Binding SID: 865
  Setup type: SR
  Bandwidth: requested 0, used 0
  LSP object:
    PLSP-ID 0x807D0, flags: D:0 S:0 R:0 A:1 O:2
  Metric type: IGP, Accumulated Metric 0
  ERO:
    SID[0]: Adj, Label 2377, NAI: local 101.102.6.1 remote 10.0.0.10
    SID[1]: Unspecified, Label 17, NAI: n/a
    SID[2]: Unspecified, Label 20, NAI: n/a

```

#### 10. Verify ODN LSP status on the PCE server.

```

Device# show pce lsp summary
PCE's LSP database summary:
-----
All peers:
Number of LSPs: 1
Operational: Up: 1 Down: 0
Admin state: Up: 1 Down: 0
Setup type: RSVP: 0 Segment routing: 1
Peer 10.0.0.4:
Number of LSPs: 1
Operational: Up: 1 Down: 0
Admin state: Up: 1 Down: 0
Setup type: RSVP: 0 Segment routing: 1

```

#### 11. Verify detailed LSP information on the PCE server.

```

Device# show pce lsp detail
PCE's tunnel database:
-----
PCC 10.0.0.4:
Tunnel Name: Tunnel2000_w
LSPs:
LSP[0]:
  source 10.0.0.4, destination 192.168.0.1, tunnel ID 2000, LSP ID 48
  State: Admin up, Operation active
  Binding SID: 872
  PCEP information:
    plsp-id 526288, flags: D:1 S:0 R:0 A:1 O:2
  Reported path:
    Metric type: IGP, Accumulated Metric 0
    SID[0]: Adj, Label 885, Address: local 10.0.0.9 remote 10.0.0.10
    SID[1]: Unknown, Label 17,
    SID[2]: Unknown, Label 20,
  Computed path:
    Computed Time: Tue Dec 20 13:12:57 2016 (00:11:53 ago)
    Metric type: IGP, Accumulated Metric 30
    SID[0]: Adj, Label 885, Address: local 10.0.0.9 remote 10.0.0.10
    SID[1]: Adj, Label 17, Address: local 10.0.0.12 remote 10.0.0.13

```

```
SID[2]: Adj, Label 20, Address: local 10.0.0.14 remote 10.0.0.14
Recorded path:
None
```

12. Shut down the interface that is connected to VRF SR so that the prefix is no longer advertised by MP-BGP.

```
Device# interface GigabitEthernet0/2
Device(config-if)# shut
```

13. Verify that VRF prefix is no longer advertised to R1 (headend) via R6 (tailend).

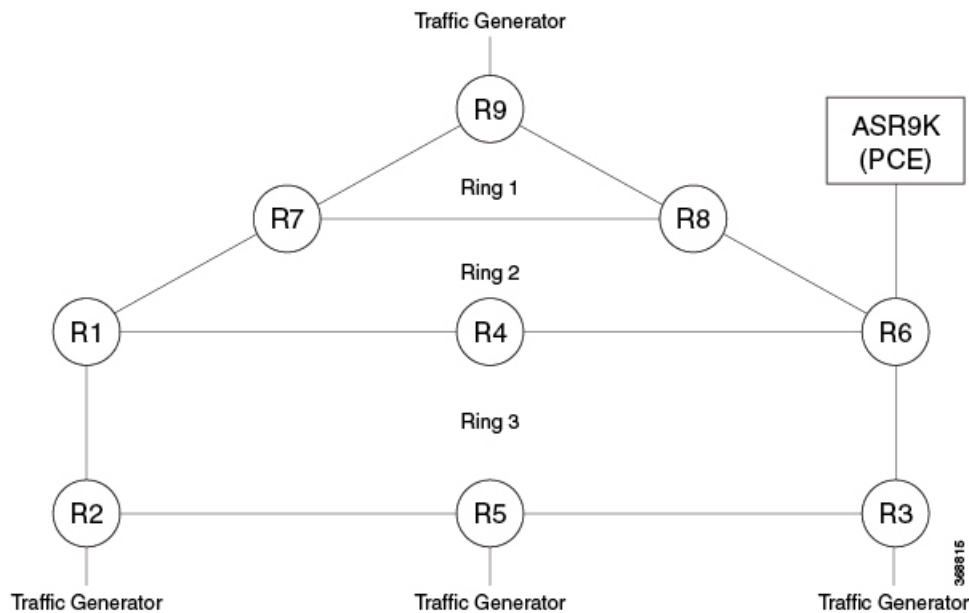
```
Device# show ip route vrf sr
Routing Table: sr
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
        D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
        ia - IS-IS inter area, * - candidate default, U - per-user static route
        o - ODR, P - periodic downloaded static route, H - NHRP, l - LISP
        a - application route
        + - replicated route, % - next hop override, p - overrides from PfR
Gateway of last resort is not set
  10.0.0.6/8 is variably subnetted, 2 subnets, 2 masks
C   10.0.0.7/24 is directly connected, GigabitEthernet0/3/1
L   10.0.0.8/32 is directly connected, GigabitEthernet0/3/1
```

14. Verify that no ODN auto-tunnel exists.

```
Device# show mpls traffic-eng tunnels
P2P TUNNELS/LSPs:
P2MP TUNNELS:
P2MP SUB-LSPS:
```

## Configuring Color Extended Community, Affinity Constraint, and Disjointness Constraint

Consider the following topology:



## Configuring Color Extended Community

### SR-TE Policy Headend Configuration on Node R3

```
segment-routing traffic-eng
on-demand color 100
authorize restrict
  ipv4 prefix-list R9350_BGP_INTER_DOMAIN
candidate-paths
  preference 1
  constraints
    segments
      dataplane mpls
  !
!
dynamic
  pcep
  !
  !
  !
!
pcc
  pce address <pce loopback ip>source-address <pcc loopback ip>
  !
```

### SR-TE Policy Tailend Configuration on Node R9

```
route-map R9_R3_R5_R2_BGP_INTER_DOMAIN permit 10
match ip address prefix-list R9350_BGP_INTER_DOMAIN
set extcommunity color 100 ----- Extended Color community configuration
route-map R9_R3_R5_R2_BGP_INTER_DOMAIN permit 20
ip prefix-list R9350_BGP_INTER_DOMAIN seq 35 permit 50.0.0.0/11 le 32
router bgp 1
address-family vpnv4
  neighbor 201.201.201.201 activate
  neighbor 201.201.201.201 send-community both
  neighbor 201.201.201.201 route-map R9_R3_R5_R2_BGP_INTER_DOMAIN out
  neighbor 206.206.206.206 activate
```

In the SR-TE ODN color template, to select the metric type, choose either **igp** or **te**:

## Configuring Affinity Constraint

## Configuring Disjointness Constraint

Segment Routing Configuration Guide, Cisco IOS XE 17 (Cisco NCS 4200 Series)

```

disjointness type node ----- Disjointness Type configuration
    source 1.0.0.0
    !
    !
    dynamic
    pcep
    !
    !
    !
    !
    pcc
    pce address <pce loopback ip> source-address <pcc loopback ip>
    !
    affinity-map
    name 1 bit-position 1

```

## Verifying SR-TE ODN Color Extended Community, Affinity Constraint, and Disjointness Constraint

### SR-TE Policy Name: 209.209.209.209|100

```

Router# show segment-routing traffic-eng policy name 209.209.209.209|100
Name: 209.209.209.209|100 (Color: 100 End-point: 209.209.209.209)
Status:
  Admin: up, Operational: up for 51:34:38 (since 01-07 06:19:08.040)- Policy state is
UP
Candidate-paths:
  Preference 1:
    Constraints:
      Affinity:
        include-any: ----- Affinity Type
        1 ----- Affinity Name
      Disjointness information:
        Group ID: 1, Source: 1.0.0.0
        Type: Node Disjointness ----- Disjointness Type
        Dynamic (pce 12.12.12.12) (active) ----- PCE Computed Candidate-path
        Weight: 0, Metric Type: TE ----- Metric Type
        Metric Type: TE, Path Accumulated Metric: 53 - Total IGP Metric from Source to
Destination
        18010 [Prefix-SID, 202.202.202.202] -----|
        18007 [Prefix-SID, 211.211.211.211] -----|
        18002 [Prefix-SID, 207.207.207.207] ----- This Segment List should
follow Affinity path
        21 [Adjacency-SID, 10.10.20.2 - 10.10.20.1] -----|
Attributes:
  Binding SID: 87 ----- Binding SID Allocated
  Allocation mode: dynamic
  State: Programmed
Auto-policy info:
  Creator: BGP SR Policy Client
  IPv6 caps enable: yes

```

### To view detailed information about SR-TE Policy 209.209.209.209|100

```

Router# show segment-routing traffic-eng policy name 209.209.209.209|100 detail
Name: 209.209.209.209|100 (Color: 100 End-point: 209.209.209.209)
Status:
  Admin: up, Operational: up for 00:04:19 (since 01-10 06:20:57.810)
Candidate-paths:
  Preference 1:

```



```

Constraints:
  Affinity:
    include-any:
      1

Disjointness information:
  Group ID: 1, Source: 1.0.0.0
  Type: Node Disjointness
  Dynamic (pce 12.12.12.12) (active)
  Weight: 0, Metric Type: TE
  Metric Type: TE, Path Accumulated Metric: 53
  18010 [Prefix-SID, 202.202.202.202]
  18007 [Prefix-SID, 211.211.211.211]
  18002 [Prefix-SID, 207.207.207.207]
  21 [Adjacency-SID, 10.10.20.2 - 10.10.20.1]

Attributes:
  Binding SID: 87
  Allocation mode: dynamic
  State: Programmed
  Auto-policy info:
    Creator: BGP SR Policy Client
    IPv6 caps enable: yes
  Forwarding-ID: 65711 (0x44) ----- This FWD-ID is used for forwarding traffic
  Stats:
    Packets: 8893      Bytes: 852848 ----- This counter indicates traffic flowing
through this SRTE policy

Event history: --- This indicates event happened with this SRTE Policy

      Timestamp                Client                Event type                Context:
Value
-----
-----
01-06 05:59:26.096            BGP SR Policy Cl      Policy created            Name:
209.209.209.209|100
01-06 05:59:26.096            BGP SR Policy Cl      Set colour                Colour: 100
01-06 05:59:26.096            BGP SR Policy Cl      Set end point            End-point:
209.209.209.209
01-06 05:59:26.096            BGP SR Policy Cl      Set dynamic pce          Path option:
dynamic pce
01-06 05:59:26.480            FH Resolution          Policy state UP           Status:
PATH RESOLVED
01-06 05:59:40.424            FH Resolution          REOPT triggered          Status:
REOPTIMIZED
01-06 05:59:49.249            FH Resolution          REOPT triggered          Status:
REOPTIMIZED
01-06 05:59:56.469            FH Resolution          REOPT triggered          Status:
REOPTIMIZED
01-07 05:15:19.918            FH Resolution          Policy state DOWN        Status:
PATH NOT RESOLVED
01-07 06:15:55.739            FH Resolution          Policy state UP          Status:
PATH RESOLVED
01-07 06:16:08.552            FH Resolution          REOPT triggered          Status:
REOPTIMIZED
01-07 06:19:08.040            FH Resolution          Policy state DOWN        Status:
PATH NOT RESOLVED
01-10 06:20:57.810            FH Resolution          Policy state UP          Status:
PATH RESOLVED
01-10 06:21:05.211            FH Resolution          REOPT triggered          Status:
REOPTIMIZED
01-10 06:21:08.036            FH Resolution          REOPT triggered          Status:
REOPTIMIZED

```

```
01-10 06:21:10.073          FH Resolution          REOPT triggered          Status:
REOPTIMIZED
```

To check if the Affinity constraint is working, shut down any of the interfaces falling under the Affinity-defined path. If the constraint works, the SR-TE policy goes down instead of taking the another path (if available) to reach to the destination.

To check if the disjointness constraint is working, check the SR-TE policy information given by the PCE, which consists of Segment IDs used for the computed path from source to destination.

Disjointness constraint works, if the Segment IDs of both the SR-TE policies are different. For example:

```
SRTE Policy 1:          SRTE Policy 2:

      SID[0]: Node, Label 16002, NAI: 207.207.207.207          SID[0]: Node, Label
16003, NAI: 208.208.208.208
      SID[1]: Node, Label 16004, NAI: 201.201.201.201          SID[1]: Node, Label 16006,
NAI: 206.206.206.206
      SID[2]: Node, Label 16011, NAI: 205.205.205.205          SID[2]: Node, Label 16011,
NAI: 205.205.205.205
```



**Note** SID[2] of policies 1 and 2 is the same since destination of both the SR-TE policies is the same.

To view the SR-TE policy and Affinity constraint in the PCE:

```
RP/0/RSP0/CPU0:ASR9K# show pce lsp pcc ipv4 213.213.213.213 private
```

```
Thu Jan 10 00:11:52.983 UTC
```

```
PCE's tunnel database:
```

```
-----
PCC 213.213.213.213:
```

```
Tunnel Name: 209.209.209.209|100
```

```
LSPs:
```

```
LSP[0]:
```

```
source 203.203.203.203, destination 209.209.209.209, tunnel ID 177, LSP ID 0
```

```
State: Admin up, Operation ---- SRTE Policy is up
```

```
Setup type: Segment Routing
```

```
Binding SID: 87
```

```
Maximum SID Depth: 4
```

```
Absolute Metric Margin: 0
```

```
Relative Metric Margin: 0%
```

```
Affinity: exclude-any 0x0 include-any 0x2 include-all 0x0 ---- This indicates Affinity
taken into account by PCE
```

```
PCEP information:
```

```
PLSP-ID 0x800b1, flags: D:1 S:0 R:0 A:1 O:2 C:0
```

```
LSP Role: Disjoint LSP
```

```
State-sync PCE: None
```

```
PCC: 213.213.213.213
```

```
LSP is subdelegated to: None
```

```
Reported path:
```

```
Metric type: TE, Accumulated Metric 53
```

```
SID[0]: Node, Label 18010, Address 202.202.202.202
```

```
SID[1]: Node, Label 18007, Address 211.211.211.211
```

```
SID[2]: Node, Label 18002, Address 207.207.207.207
```

```
SID[3]: Adj, Label 21, Address: local 10.10.20.2 remote 10.10.20.1
```

```
Computed path: (Local PCE)
```

```
Computed Time: Thu Jan 10 00:09:36 UTC 2019 (00:02:17 ago)
```

```
Metric type: TE, Accumulated Metric 53
```

```

SID[0]: Node, Label 18010, Address 202.202.202.202
SID[1]: Node, Label 18007, Address 211.211.211.211
SID[2]: Node, Label 18002, Address 207.207.207.207
SID[3]: Adj, Label 21, Address: local 10.10.20.2 remote 10.10.20.1

```

Recorded path:

None

Disjoint Group Information:

Type Node-Disjoint, Group 1, Sub-Group 1.0.0.0

Event history (latest first):

Time	Event
Thu Jan 10 00:09:37 UTC 2019	Report from 213.213.213.213 (LSP owner) Symbolic-name: 209.209.209.209 100, LSP-ID: 0, Source: 203.203.203.203 Destination: 209.209.209.209, D:1, R:0, A:1 O:2, Sig.BW: 0, Act.BW: 0 Reported Path: (Metric 53) Label 18010, Address 202.202.202.202 Label 18007, Address 211.211.211.211 Label 18002, Address 207.207.207.207 Label 21, Address: local 10.10.20.2 remote 10.10.20.1 Chng:0, AssoChng:0
Thu Jan 10 00:09:36 UTC 2019	Update to 213.213.213.213 (PCC) Symbolic-name: 209.209.209.209 100, LSP-ID: 0, D:1 Path: (Metric 53) Label 18010, Address 202.202.202.202 Label 18007, Address 211.211.211.211 Label 18002, Address 207.207.207.207 Label 21, Address: local 10.10.20.2 remote 10.10.20.1
Thu Jan 10 00:09:36 UTC 2019	Path Computation (Disjoint LSP) Symbolic-name: 209.209.209.209 100, LSP-ID: 0, D:1 Source: 203.203.203.203 Destination: 209.209.209.209 Status: Disjoint Path Success
Wed Jan 09 23:54:42 UTC 2019	Update to 213.213.213.213 (PCC) Symbolic-name: 209.209.209.209 100, LSP-ID: 0, D:1 Path: (Metric 53) Label 18007, Address 211.211.211.211 Label 18002, Address 207.207.207.207 Label 21, Address: local 10.10.20.2 remote 10.10.20.1
Wed Jan 09 23:54:42 UTC 2019	Path Computation (Disjoint LSP) Symbolic-name: 209.209.209.209 100, LSP-ID: 0, D:1 Source: 203.203.203.203 Destination: 209.209.209.209 Status: Fallback Node to Shortest Path Computed Path: (Metric 53) Label 18007, Address 211.211.211.211 Label 18002, Address 207.207.207.207 Label 21, Address: local 10.10.20.2 remote 10.10.20.1
Wed Jan 09 23:54:21 UTC 2019	Path Computation (Disjoint LSP) Symbolic-name: 209.209.209.209 100, LSP-ID: 0, D:1 Source: 203.203.203.203 Destination: 209.209.209.209 Status: Disjoint Path Success
Computed Path: (Metric 53)	Label 18010, Address 202.202.202.202 Label 18007, Address 211.211.211.211 Label 18002, Address 207.207.207.207 Label 21, Address: local 10.10.20.2 remote 10.10.20.1
Computed Path: (Metric 53)	Label 18010, Address 202.202.202.202 Label 18007, Address 211.211.211.211

```

Thu Jan 10 00:09:05 UTC 2019
Label 18002, Address 207.207.207.207
Label 21, Address: local 10.10.20.2 remote 10.10.20.1
Path Computation (Disjoint LSP)
Symbolic-name: 209.209.209.209|100, LSP-ID: 0, D:1
Source: 203.203.203.203 Destination: 209.209.209.209
Status: Fallback Node to Shortest Path
Computed Path: (Metric 53)
Label 18007, Address 211.211.211.211
Label 18002, Address 207.207.207.207
Label 21, Address: local 10.10.20.2 remote 10.10.20.1
Wed Jan 09 23:54:42 UTC 2019
Report from 213.213.213.213 (LSP owner)
Symbolic-name: 209.209.209.209|100, LSP-ID: 0,
Source: 203.203.203.203 Destination: 209.209.209.209,
D:1, R:0, A:1 O:2, Sig.BW: 0, Act.BW: 0
Reported Path: (Metric 53)
Label 18007, Address 211.211.211.211
Label 18002, Address 207.207.207.207
Label 21, Address: local 10.10.20.2 remote 10.10.20.1
Chng:0, AssoChng:0

RP/0/RSP0/CPU0:ASR9K#

To view disjointness between policies 1 and 2:

RP/0/RSP0/CPU0:ASR9K# show pce association type link group-id 3

Wed Aug 29 05:56:52.228 UTC
PCE's association database:
-----
Association: Type Link-Disjoint, Group 3, Sub-Group 1.0.0.0, Not Strict
Associated LSPs:
  LSP[0]:
    PCC 213.213.213.213, tunnel name 209.209.209.209|104, PLSP ID 524460, tunnel ID 172,
LSP ID 0, Configured on PCC
  LSP[1]:
    PCC 213.213.213.213, tunnel name 209.209.209.209|105, PLSP ID 524461, tunnel ID 173,
LSP ID 0, Configured on PCC
Status: Satisfied ----- This indicates that Disjointness between SRTE Policies
is working
RP/0/RSP0/CPU0:ASR9K#

```

## Troubleshooting the SR-TE ODN Color Extended Community, Affinity Constraint, and Disjointness Constraint

If SR-TE policy is down, check the status of the SR-TE Policy under the SR-TE policy information

```

Router# show segment-routing traffic-eng policy name 209.209.209.209|100
Name: 209.209.209.209|106 (Color: 106 End-point: 209.209.209.209)
Status:
  Admin: up, Operational: down for 00:00:18 (since 01-10 13:06:42.142)
Candidate-paths:
  Preference 1:
    Constraints:
      Affinity:
        include-any:
          1
      Dynamic (pce) (inactive)
      Weight: 0, Metric Type: IGP
Attributes:
  Binding SID: 269
  Allocation mode: dynamic
  State: Programmed

```

```

Auto-policy info:
  Creator: BGP SR Policy Client
  IPv6 caps enable: yes
Router#

```



**Note** The possible reasons for the policy being down are:

- Connection to PCE is down.
- Max SID depth is exceeded.
- An interface falling under Affinity-defined path from source to destination has been shut down.

To check the SR-TE policy status on the PCE:

```

RP/0/RSP0/CPU0:ASR9K#show pce lsp pcc ipv4 213.213.213.213 private
Thu Jan 10 00:11:52.983 UTC
PCE's tunnel database:
-----
PCC 213.213.213.213:
Tunnel Name: 209.209.209.209|100
LSPs:
LSP[0]:
  source 203.203.203.203, destination 209.209.209.209, tunnel ID 177, LSP ID 0
  State: Admin up, Operation active      ----- SRTE Policy is up
  Setup type: Segment Routing
  Binding SID: 87

```

```

Maximum SID Depth: 4
Absolute Metric Margin: 0
Relative Metric Margin: 0%

```

Affinity: exclude-any 0x0 include-any 0x2 include-all 0x0 --- This indicates Affinity is taken into account by the PCE

PCE is aware of the network topology. This information is used for path computation using the following command. This information is also used to determine if nodes and links are present and have the expected attributes (IGP/TE admin weights, SIDs and so on).

```

RP/0/RSP0/CPU0:ASR9K# show pce ipv4 topology
Tue Jan 15 01:36:20.298 UTC
PCE's topology database - detail:
-----
Node 1
  TE router ID: 207.207.207.207
  Host name: 920-R7
  ISIS system ID: 0000.0000.0207 level-1 ASN: 1
  ISIS system ID: 0000.0000.0207 level-2 ASN: 1
  Prefix SID:
  ISIS system ID: 0000.0000.0207 level-1 ASN: 1 domain ID: 0
    Prefix 207.207.207.207, label 16002 (regular), flags: N
  ISIS system ID: 0000.0000.0207 level-1 ASN: 1 domain ID: 0
    Prefix 207.207.207.207, label 18002 (strict), flags: N
  ISIS system ID: 0000.0000.0207 level-2 ASN: 1 domain ID: 0
    Prefix 207.207.207.207, label 16002 (regular), flags: N
  ISIS system ID: 0000.0000.0207 level-2 ASN: 1 domain ID: 0
    Prefix 207.207.207.207, label 18002 (strict), flags: N
SRGB INFO:

```

```

isis system ID: 0000.0000.0207 level-1 ASN: 1
  SRGB Start: 16000 Size: 8000
isis system ID: 0000.0000.0207 level-2 ASN: 1
  SRGB Start: 16000 Size: 8000
Link[0]: local address 10.10.21.1, remote address 10.10.21.2
  Local node:
    isis system ID: 0000.0000.0207 level-1 ASN: 1
  Remote node:
    TE router ID: 208.208.208.208
    Host name: 920-R8
    isis system ID: 0000.0000.0208 level-1 ASN: 1
    Metric: IGP 10, TE 10, Latency 10
    Bandwidth: Total 1250000000 Bps, Reservable 0 Bps
    Admin-groups: 0x00000000
    Adj SID: 16 (unprotected) 17 (protected)
Link[1]: local address 10.10.21.1, remote address 10.10.21.2
  Local node:
isis system ID: 0000.0000.0207 level-2 ASN: 1
  Remote node:
    TE router ID: 208.208.208.208
    Host name: 920-R8
    isis system ID: 0000.0000.0208 level-2 ASN: 1
    Metric: IGP 10, TE 10, Latency 10
    Bandwidth: Total 1250000000 Bps, Reservable 0 Bps
    Admin-groups: 0x00000000
    Adj SID: 18 (unprotected) 19 (protected)
Link[2]: local address 10.10.20.2, remote address 10.10.20.1
  Local node:
    isis system ID: 0000.0000.0207 level-2 ASN: 1
  Remote node:
    TE router ID: 209.209.209.209
    Host name: 920-R9
    isis system ID: 0000.0000.0209 level-2 ASN: 1
    Metric: IGP 40, TE 40, Latency 40
    Bandwidth: Total 1250000000 Bps, Reservable 0 Bps
    Admin-groups: 0x00000052
    Adj SID: 20 (unprotected) 22 (protected)
    SRLG Values: 25
Node 2
  TE router ID: 209.209.209.209
  Host name: 920-R9
  isis system ID: 0000.0000.0209 level-1 ASN: 1
  isis system ID: 0000.0000.0209 level-2 ASN: 1
  Prefix SID:
    isis system ID: 0000.0000.0209 level-1 ASN: 1 domain ID: 0
      Prefix 209.209.209.209, label 16001 (regular), flags: N
    isis system ID: 0000.0000.0209 level-1 ASN: 1 domain ID: 0
      Prefix 209.209.209.209, label 18001 (strict), flags: N
    isis system ID: 0000.0000.0209 level-2 ASN: 1 domain ID: 0
      Prefix 209.209.209.209, label 16001 (regular), flags: N
    isis system ID: 0000.0000.0209 level-2 ASN: 1 domain ID: 0
      Prefix 209.209.209.209, label 18001 (strict), flags: N
  SRGB INFO:
    isis system ID: 0000.0000.0209 level-1 ASN: 1
      SRGB Start: 16000 Size: 8000
    isis system ID: 0000.0000.0209 level-2 ASN: 1
      SRGB Start: 16000 Size: 8000
Link[0]: local address 10.10.20.1, remote address 10.10.20.2
  Local node:
    isis system ID: 0000.0000.0209 level-2 ASN: 1
  Remote node:
    TE router ID: 207.207.207.207
    Host name: 920-R7
    isis system ID: 0000.0000.0207 level-2 ASN: 1

```

```
Metric: IGP 40, TE 40, Latency 40
Bandwidth: Total 1250000000 Bps, Reservable 0 Bps
Admin-groups: 0x00000052
Adj SID: 1980 (unprotected) 1981 (protected)
Link[1]: local address 10.10.22.1, remote address 10.10.22.2
Local node:
  ISIS system ID: 0000.0000.0209 level-2 ASN: 1
Remote node:
  TE router ID: 208.208.208.208
  Host name: 920-R8
  ISIS system ID: 0000.0000.0208 level-2 ASN: 1
Metric: IGP 10, TE 50, Latency 50
Bandwidth: Total 1250000000 Bps, Reservable 0 Bps
Admin-groups: 0x0000002C
Adj SID: 1971 (unprotected) 1972 (protected)
```

RP/0/RSP0/CPU0:ASR9K#

Further troubleshooting tips:

- Enable the following debug commands on the PCCs:
  - debug segment-routing traffic-eng path
  - debug segment-routing traffic-eng pcalc
  - debug segment-routing traffic-eng policy
  - debug segment-routing traffic-eng topology
  - debug segment-routing traffic-eng ha
- Enable the following debug commands on the PCE:
  - debug pce pcep
  - debug pce cspf
  - debug pce cspf-internal
  - debug pce error
  - debug pce path







## CHAPTER 10

# Segment Routing OAM Support

Effective Cisco IOS XE Everest 16.6.1, you can verify that the operation with Segment Routing OAM label switched protocol ping and traceroute (SR OAM LSPV).



### Note

OAM Support is available only on the Cisco RSP2 Module.

- [Restrictions for Segment Routing OAM Support, on page 115](#)
- [Information About Segment Routing-OAM Support , on page 115](#)
- [How to Diagnose Segment Routing with LSP Ping and Trace Route Nil FEC Target, on page 116](#)
- [Example: LSP Ping Nil\\_FEC Target Support, on page 117](#)

## Restrictions for Segment Routing OAM Support

- Ping and traceroute do not display proper output over SR-TE tunnels using verbatim path option.
- Ping and traceroute are unsupported with SR-TE Static auto tunnel, BGP Dynamic TE, and On-demand next hop auto tunnels.

## Information About Segment Routing-OAM Support

### Segment Routing-OAM Support

The Segment Routing-OAM Support feature provides support for Nil-FEC LSP Ping/Trace functionality. Nil-FEC LSP Ping/Trace functionality support Segment Routing and MPLS Static. It also acts as an additional diagnostic tool for all other LSP types. This feature allows operators to provide the ability to freely test any label stack by allowing them to specify the following:

- Label stack
- Outgoing interface
- Nexthop address

In the case of segment routing, each segment nodal label and adjacent label along the routing path is put into the label stack of an echo request message from initiator Label Switch Router (LSR); MPLS data plane forward this packet to the label stack target, and the label stack target reply the echo message back.

## LSP Ping Operation for Nil FEC target

The LSP Ping/Traceroute is used in identifying LSP breakages. The nil-fec target type can be used to test the connectivity for a known label stack. Follow the existing LSP ping procedure (for more information, refer MPLS LSP Ping/Traceroute), with the following modifications:

- Build the echo request packet with the given label stack;
- Append explicit null label at the bottom of the label stack;
- Build echo request FTS TLV with target FEC Nil FEC and label value set to the bottom label of the label stack, which is explicit-null.

## How to Diagnose Segment Routing with LSP Ping and Trace Route Nil FEC Target

### Use LSP Ping for Nil FEC Target

The Nil-FEC LSP ping and traceroute operation are simply extension of regular MPLS ping and trace route. nil-fec labels <label, label...> is added to the ping mpls command. This command sends an echo request message with MPLS label stack as specified, and add another explicit null at bottom of the stack.

```
ping mpls
  {{ipv4 <target>/<mask> [fec-type {bgp | generic | ldp}] |
   {pseudowire <peer addr> <vc-id> [segment<segment-number>]} |
   {traffic-eng {<tunnel interface> |
    {p2p <sender> <endpoint>
     <tun-id> <ex-tun-id> <lspid>} |
    {p2mp <p2mp-id> <sender>
     <tun-id><ex-tun-id> <lspid>}}} |
   {tp <tunnel-tp interface> lsp <working | protect | active>}}} |
  {mldp {p2mp | mp2mp}
  {{ipv4 <source> <group>} |
  {ipv6 <source> <group>} |
  {vpnv4 <rd> <source> <group>} |
  {vpnv6 <rd> <source> <group>} |
  {hex <opaque type> <opaque value>} |
  {mdt <vpnid> <mdt-num>} |
  {static-id <lsp-identifier>} |
  {global-id <global-identifier>}}} |
  {nil-fec {labels <comma separated labels>}}
  [repeat <count>]
  [size <size> | sweep <min_size> <max_size> <increment>]
  [timeout <seconds>]
  [interval <milliseconds>]
  [destination <addr_start> [<addr_end> [<addr_incr_mask> | <addr_incr>]]]
  [source <addr>]
  [exp <exp-value>]
  [pad <pattern>]
  [ttl <ttl>]
  [reply [mode [ipv4 | router-alert | no-reply]]]
  [dscp <dscp-bits>]
  [pad-tlv]
  [verbose]
  [force-explicit-null]
  [force-disposition ra-label]
```

```
[output {interface <tx-interface>} [nexthop <nexthop ip addr>]]
[ {dsmapi | ddmapi [l2ecmpi]} [hashkey {none | {ipv4 | ipv4-label-set {bitmap <bitmap_size>}}}]
[flags {fec | ttl}]
[segment {all | <lower-segment-number> [upper-segment-number]}]
[jitter <milliseconds>]
[responder-id <ip addr> | egress <ipaddr>]
  NOTE: responder-id for mldp/ egress for p2mp te
[entropy-label <label-value>]
```

For more information, see [ping mpls](#).

## Using LSP Traceroute for Nil FEC Target

```
trace mpls
  {{ipv4 <target>/<mask> [fec-type {bgp | generic | ldp}]}} |
  {traffic-eng {<tunnel interface> |
    {p2p <sender> <endpoint>
      <tun-id> <ex-tun-id> <lspid>} |
    {p2mp <p2mp-id> <sender>
      <tun-id><ex-tun-id> <lspid>}}} |
  {mldp {p2mp | mp2mp}}
  {{ip <source> <group>} |
  {vpn <rd> [<source> <group>]} |
  {hex <opaque type> <opaque value>} |
  {mdt <vpnid> <mdt-num>} |
  {static-id <lsp-identifier>} |
  {global-id <global-identifier>}}}
  {pseudowire <next-pe-address> <pwid> [segment
    <segment-number>/<lower-segment-number> [<upper-segment-limit>]]} |
  {tp <tunnel-tp interface> lsp <working | protect | active>}} |
  {nil-fec {labels <comma separated labels>}}
  [timeout <seconds>]
  [destination <addr_start> [<addr_end> [<addr_incr_mask> | <addr_incr>]]]
  [source <addr> ]
  [exp <exp-value>]
  [ttl <ttl-max>]
  [reply [mode {ipv4 | router-alert | no-reply}]]
  [dscp <dscp-bits>]
  [pad-tlv]
  [force-explicit-null]
  [output {interface <tx-interface>} [nexthop <nexthop ip addr>]]
  [flags {fec | ttl}]
  [entropy-label <label-value>]
[hashkey ipv4 | ipv4-label-set {bitmap <bitmap_size>}]
```

For more information, see [traceroute mpls](#).

## Example: LSP Ping Nil\_FEC Target Support

Node loopback IP address:	1.1.1.3	1.1.1.4	1.1.1.5
	1.1.1.7		
Node label:		16004	16005
	16007		
Nodes:	Arizona	Utah	Wyoming
----- Texas			
Interface:	Eth1/0	Eth1/0	Interface IP address:
			30.1.1.3
	30.1.1.4		

## Example: LSP Ping Nil\_FEC Target Support

```

Device# show mpls forwarding-table
Local   Outgoing Prefix          Bytes Label  Outgoing Next Hop
Label   Label   or Tunnel Id      Switched    interface
16      0       Et1/0            30.1.1.3    \
Pop Label 3333.3333.0000-Et1/0-30.1.1.3 \
17      0       Et1/1            90.1.1.5    \
Pop Label 5555.5555.5555-Et1/1-90.1.1.5 \
18      0       Et0/2            102.102.102.2 \
Pop Label 3333.3333.0253-Et0/2-102.102.102.2 \
19      Pop Label 9.9.9.4/32      0           Et0/2      102.102.102.2
20      Pop Label 1.1.1.5/32      0           Et1/1      90.1.1.5
21      Pop Label 1.1.1.3/32      0           Et1/0      30.1.1.3
22      Pop Label 16.16.16.16/32 0           Et1/0      30.1.1.3
23      Pop Label 16.16.16.17/32 0           Et1/0      30.1.1.3
24      Pop Label 17.17.17.17/32 0           Et1/0      30.1.1.3
25      20      9.9.9.3/32      0           Et1/0      30.1.1.3
26      21      1.1.1.6/32      0           Et1/0      30.1.1.3
27      24      1.1.1.2/32      0           Et1/0      30.1.1.3
28      1.1.1.2/32 0           Et1/1      90.1.1.5
28      18      1.1.1.7/32      0           Et1/1      90.1.1.5
29      27      9.9.9.7/32      0           Et1/1      90.1.1.5
30      Pop Label 55.1.1.0/24      0           Et1/1      90.1.1.5
31      Pop Label 19.1.1.0/24      0           Et1/0      30.1.1.3
Local   Outgoing Prefix          Bytes Label  Outgoing Next Hop
Label   Label   or Tunnel Id      Switched    interface
32      Pop Label 100.1.1.0/24      0           Et1/0      30.1.1.3
33      Pop Label 100.100.100.0/24 0           Et1/0      30.1.1.3
34      Pop Label 110.1.1.0/24      0           Et1/0      30.1.1.3
35      28      10.1.1.0/24      0           Et1/0      30.1.1.3
36      29      101.101.101.0/24 0           Et1/0      30.1.1.3
37      29      65.1.1.0/24      0           Et1/1      90.1.1.5
38      33      104.104.104.0/24 0           Et1/0      30.1.1.3
39      104.104.104.0/24 0           Et1/1      90.1.1.5
39      30      103.103.103.0/24 0           Et1/1      90.1.1.5
16005   Pop Label 1.1.1.5/32      1782        Et1/1      90.1.1.5
16006   16006   1.1.1.6/32      0           Et1/0      30.1.1.3
16007   16007   1.1.1.7/32      0           Et1/1      90.1.1.5
16017   16017   17.17.17.17/32 0           Et1/0      30.1.1.3
16250   16250   9.9.9.3/32      0           Et1/0      30.1.1.3
16252   16252   9.9.9.7/32      0           Et1/1      90.1.1.5
16253   Pop Label 9.9.9.4/32      0           Et0/2      102.102.102.2
17000   17000   16.16.16.16/32 0           Et1/0      30.1.1.3
17002   17002   1.1.1.2/32      0           Et1/0      30.1.1.3
17002   1.1.1.2/32 0           Et1/1      90.1.1.5
SR231-arizona#ping mpls nil-fec labels 16005,16007 output interface ethernet 1/0 nexthop
30.1.1.4 repeat 1
Sending 1, 72-byte MPLS Echos with Nil FEC labels 16005,16007, timeout is 2 seconds, send
interval
is 0 msec:
Codes: '!' - success, 'Q' - request not sent, '.' - timeout,
'L' - labeled output interface, 'B' - unlabeled output interface,
'D' - DS Map mismatch, 'F' - no FEC mapping, 'f' - FEC mismatch,
'M' - malformed request, 'm' - unsupported tlvs, 'N' - no label entry,
'P' - no rx intf label prot, 'p' - premature termination of LSP,
'R' - transit router, 'I' - unknown upstream index,
'l' - Label switched with FEC change, 'd' - see DDMAP for return code,
'X' - unknown return code, 'x' - return code 0
Type escape sequence to abort.
!
Success rate is 100 percent (1/1), round-trip min/avg/max = 1/1/1 ms
Total Time Elapsed 0 ms
SR231-arizona#traceroute mpls nil-fec labels 16005,16007 output interface ethernet 1/0
nexthop 30.1.1.4
Tracing MPLS Label Switched Path with Nil FEC labels 16005,16007, timeout is 2 seconds

```

Codes: '!' - success, 'Q' - request not sent, '.' - timeout,  
'L' - labeled output interface, 'B' - unlabeled output interface,  
'D' - DS Map mismatch, 'F' - no FEC mapping, 'f' - FEC mismatch,  
'M' - malformed request, 'm' - unsupported tlvs, 'N' - no label entry,  
'P' - no rx intf label prot, 'p' - premature termination of LSP,  
'R' - transit router, 'I' - unknown upstream index,  
'l' - Label switched with FEC change, 'd' - see DDMAP for return code,  
'X' - unknown return code, 'x' - return code 0  
Type escape sequence to abort.  
0 30.1.1.3 MRU 1500 [Labels: 16005/16007/explicit-null Exp: 0/0/0]  
L 1 30.1.1.4 MRU 1500 [Labels: implicit-null/16007/explicit-null Exp: 0/0/0] 1 ms  
L 2 90.1.1.5 MRU 1500 [Labels: implicit-null/explicit-null Exp: 0/0] 1 ms  
! 3 55.1.1.7 1 ms





## CHAPTER 11

# Dynamic Path Computation Client

Effective Cisco IOS XE Everest 16.6.1, the Stateful Path Computation Element Protocol (PCEP) enables a router to report and delegate Label Switched Paths (LSPs) that are established using Resource Reservation Protocol (RSVP) protocol to a stateful Path Computation Element (PCE). An LSP delegated to a PCE can be updated by the PCE and a stateful PCE can initiate an LSP on a Path Computation Client (PCC).

RSVP-TE LSPs require link-state routing protocols such as OSPF or IS-IS to distribute and learn traffic engineering topology. A stateful PCE can learn the traffic engineering topology through BGP Link-State (BGP-LS) protocol. The IGP instance running on a router feeds the traffic engineering topology information into BGP which passes that information to stateful PCE.



---

**Note** Dynamic PCC is available only on the Cisco RSP2 Module.

---



---

**Note** Path Computation Element is not supported on the Cisco RSP2 Module.

---

- [Information About Dynamic PCC , on page 121](#)
- [Configuring Dynamic PCC Globally, on page 122](#)
- [Configuring Dynamic PCC on an Interface, on page 122](#)
- [Verifying Dynamic PCC, on page 122](#)

## Information About Dynamic PCC

### PCEP functions

A PCEP session is a TCP session between a PCC and a PCE with protocol messages. The PCEP functions are verified based on the PCC functions. The configuration and verification show that the request is accepted and path computation is provided based on PCReq message from the client. The passive reporting enables a router to report a tunnel instead of delegating it to a PCE. The PCE is aware of the tunnel even though it cannot modify the tunnel.

PCEP functions are useful when a network has both router-controlled and PCE delegated tunnels. The PCE is aware of both the tunnels and can make an accurate decision on path computation.

## Configuring Dynamic PCC Globally

Perform the following task to configure dynamic PCC globally:

```
enable
configure terminal
mpls traffic-eng tunnels
mpls traffic-eng pcc peer 10.0.0.1 ----..(10.0.0.1 is the PCE server address)
mpls traffic-eng pcc report-all
end
```

## Configuring Dynamic PCC on an Interface

Perform the following task to configure dynamic PCC on an interface

```
interface Tunnell
ip unnumbered Loopback0
tunnel mode mpls traffic-eng
tunnel destination 7.7.7.7
tunnel mpls traffic-eng autoroute announce
tunnel mpls traffic-eng priority 5 5
tunnel mpls traffic-eng bandwidth 200
tunnel mpls traffic-eng path-option 10 dynamic pce segment-routing
end
```

## Verifying Dynamic PCC

The following sample output is from the **show pce client peer detail** command.

```
Device# show pce client peer detail
PCC's peer database:
-----
Peer address: 1.1.1.1
  State up
  Capabilities: Stateful, Update, Segment-Routing
  PCEP has been up for: 23:44:58
  PCEP session ID: local 1, remote: 0
  Sending KA every 30 seconds
  Minimum acceptable KA interval: 20 seconds
  Peer timeout after 120 seconds
  Statistics:
Keepalive messages: rx      2798 tx      2112
Request messages:  rx         0 tx         32
Reply messages:    rx       32 tx         0
Error messages:    rx         0 tx         0
Open messages:     rx         1 tx         1
Report messages:   rx         0 tx        57
Update messages:   rx       72 tx         0
```

The following sample output is from the **show tunnel** command which shows the LSP details.



```

Device# show tunnel 1
Name: dl_t1                                     (Tunnel) Destination: 7.7.7.7
Status:
  Admin: up          Oper: up          Path: valid          Signalling: connected
  path option 10, (SEGMENT-ROUTING) (PCE) type dynamic (Basis for Setup, path weight 0)
Config Parameters:
  Bandwidth: 200      kbps (Global) Priority: 5 5 Affinity: 0x0/0xFFFF
  Metric Type: TE (default)
  Path Selection:
    Protection: any (default)
  Path-selection Tiebreaker:
    Global: not set Tunnel Specific: not set Effective: min-fill (default)
  Hop Limit: disabled
  Cost Limit: disabled
  Path-invalidation timeout: 10000 msec (default), Action: Tear
  AutoRoute: enabled LockDown: disabled Loadshare: 200 [10000000] bw-based
  auto-bw: disabled
  Fault-OAM: disabled, Wrap-Protection: disabled, Wrap-Capable: No
Active Path Option Parameters:
  State: dynamic path option 10 is active
  BandwidthOverride: disabled LockDown: disabled Verbatim: disabled
PCEP Info:
  Delegation state: Working: yes Protect: no
  Current Path Info:
    Request status: processed
    Created via PCRep message from PCE server: 1.1.1.1
  Reported paths:
    Tunnel Name: csr551_t2001
    LSPs:
      LSP[0]:
        source 2.2.2.2, destination 7.7.7.7, tunnel ID 1, LSP ID 5
        State: Admin up, Operation active
        Setup type: SR
        Bandwidth: signaled 0
        LSP object:
          PLSP-ID 0x807D1, flags: D:0 S:0 R:0 A:1 O:2
        Reported path:
          Metric type: TE, Accumulated Metric 0
History:
  Tunnel:
    Time since created: 34 minutes, 3 seconds
    Time since path change: 1 minutes, 44 seconds
    Number of LSP IDs (Tun_Instances) used: 5
    Current LSP: [ID: 5]
    Uptime: 1 minutes, 44 seconds
    Prior LSP: [ID: 3]
    ID: path option unknown
    Removal Trigger: path verification failed
  Tun_Instance: 5
Segment-Routing Path Info (isis level-1)
Segment0[Node]: 3.3.3.3, Label: 20270
Segment1[Node]: 6.6.6.6, Label: 20120
Segment2[Node]: 7.7.7.7, Label: 20210

```

The following sample output is from the **show pce client lsp detail** command.

```

Device# show pce client lsp detail
PCC's tunnel database:
-----
Tunnel Name: dl_t1
LSPs:
LSP[0]:

```

```

source 2.2.2.2, destination 7.7.7.7, tunnel ID 1, LSP ID 5
State: Admin up, Operation active
Setup type: SR
Bandwidth: signaled 0
LSP object:
  PLSP-ID 0x807D1, flags: D:0 S:0 R:0 A:1 O:2
Reported path:
  Metric type: TE, Accumulated Metric 0

```

The following sample output is from the **show pce lsp detail** command which shows the tunnel is delegated.

```

Device# show pce lsp detail
Thu Jul  7 10:24:30.836 EDT
PCE's tunnel database:
-----
PCC 102.103.2.1:
Tunnel Name: dl_t1
LSPs:
LSP[0]:
  source 2.2.2.2, destination 7.7.7.7, tunnel ID 1, LSP ID 5
  State: Admin up, Operation active
  Binding SID: 0
  PCEP information:
    plsp-id 526289, flags: D:1 S:0 R:0 A:1 O:2
  Reported path:
    Metric type: TE, Accumulated Metric 0
    SID[0]: Node, Label 20270, Address 3.3.3.3
    SID[1]: Node, Label 20120, Address 6.6.6.6
    SID[2]: Node, Label 20210, Address 7.7.7.7
  Computed path:
    Metric type: TE, Accumulated Metric 30
    SID[0]: Node, Label 20270, Address 3.3.3.3
    SID[1]: Node, Label 20120, Address 6.6.6.6
    SID[2]: Node, Label 20210, Address 7.7.7.7
  Recorded path:
    None

```

The following sample output is from the **show pce client lsp detail** command for reported tunnel.

```

Device# show pce client lsp detail
PCC's tunnel database:
-----
Tunnel Name: dl_t2
LSPs:
LSP[0]:
  source 2.2.2.2, destination 7.7.7.7, tunnel ID 2, LSP ID 1
  State: Admin up, Operation active
  Setup type: SR
  Bandwidth: signaled 0
  LSP object:
    PLSP-ID 0x807D2, flags: D:0 S:0 R:0 A:1 O:2
  Reported path:
  Metric type: TE, Accumulated Metric 30

```

The following sample output is from the **show pce lsp detail** command which shows the tunnel is not delegated.

```

Device# show pce lsp detail
Thu Jul  7 10:29:48.754 EDT
PCE's tunnel database:
-----

```

```
PCC 10.0.0.1:
Tunnel Name: dl_t2
LSPs:
LSP[0]:
  source 2.2.2.2, destination 7.7.7.7, tunnel ID 2, LSP ID 1
  State: Admin up, Operation active
  Binding SID: 0
  PCEP information:
    plsp-id 526290, flags: D:0 S:0 R:0 A:1 O:2
  Reported path:
    Metric type: TE, Accumulated Metric 30
    SID[0]: Adj, Label 74, Address: local 172.16.0.1 remote 172.16.0.2
    SID[1]: Adj, Label 63, Address: local 173.17.0.1 remote 173.17.0.2
    SID[2]: Adj, Label 67, Address: local 174.18.0.1 remote 174.18.0.2
    SID[3]: Node, Label unknownAddress 7.7.7.7
  Computed path:
    None
  Recorded path:
    None
```





## CHAPTER 12

# Segment Routing uLoop Avoidance

- [Prerequisites for Segment Routing uLoop Avoidance, on page 127](#)
- [Restrictions for Segment Routing uLoop Avoidance, on page 127](#)
- [Information About Segment Routing uLoop Avoidance, on page 127](#)
- [How to Enable Segment Routing uLoop Avoidance, on page 131](#)

## Prerequisites for Segment Routing uLoop Avoidance

- The SR: uLoop Avoidance feature will work only if the Topology-Independent Loop-Free Alternate (TI-LFA) feature is configured.

## Restrictions for Segment Routing uLoop Avoidance

- Segment routing uLoop avoidance feature does not support double fault at the same time.
- Node up/down events are not supported in OSPF uLoop avoidance.

## Information About Segment Routing uLoop Avoidance

### Microloops

When changes occur in a network topology because of the failure or restoration of a link or a network device, IP Fast Reroute enables rapid network convergence by moving traffic to precomputed backup paths until regular convergence mechanisms move traffic to a newly computed best path, also known as a post-convergence path. This network convergence may cause short microloops between two directly or indirectly connected devices in the topology. Microloops are caused when different nodes in the network calculate alternate paths at different times and independently of each other. For instance, if a node converges and sends traffic to a neighbor node, which has not converged yet, traffic may loop between the two nodes.

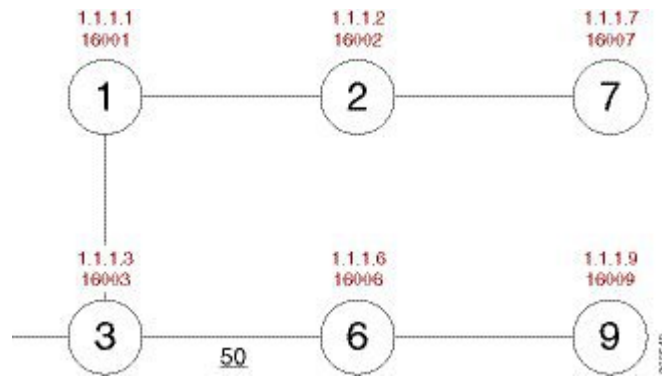
Microloops may or may not result in traffic loss. If the duration of a microloop is short, that is the network converges quickly, packets may loop for a short duration before their time-to-live (TTL) expires. Eventually, the packets get forwarded to the destination. If the duration of the microloop is long, that is one of the routers

in the network is slow to converge, packets may expire their TTL or the packet rate may exceed the bandwidth, or the packets might be out of order, and packets may get dropped.

Microloops that are formed between a failed device and its neighbors are called local uloops, whereas microloops that are formed between devices that are multiple hops away are called remote uloops. Local uloops are usually seen in networks where local loop-free alternate (LFA) path is not available. In such networks, remote LFAs provide backup paths for the network.

The information discussed above can be illustrated with the help of an example topology as shown in the following figure.

**Figure 10: Microloop Example Topology**



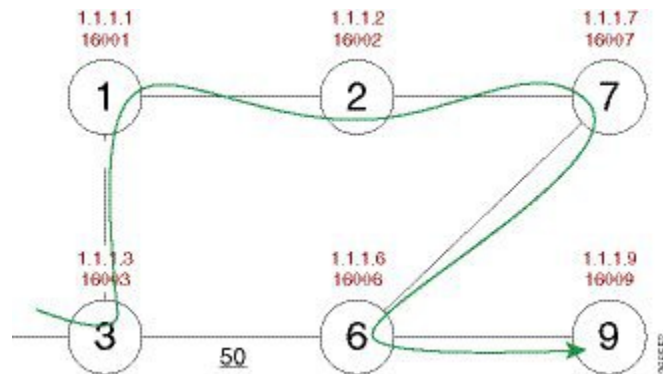
The assumptions in this example are as follows:

- The default metrics is 10 for each link except for the link between Node 3 and Node 6, which has a metric of 50. The order of convergence with SPF backoff delays on each node is as follows:
  - Node 3—50 milliseconds
  - Node 1—500 milliseconds
  - Node 2—1 second
  - Node 2—1.5 seconds

A packet sent from Node 3 to Node 9, the destination, traverses via Node 6.

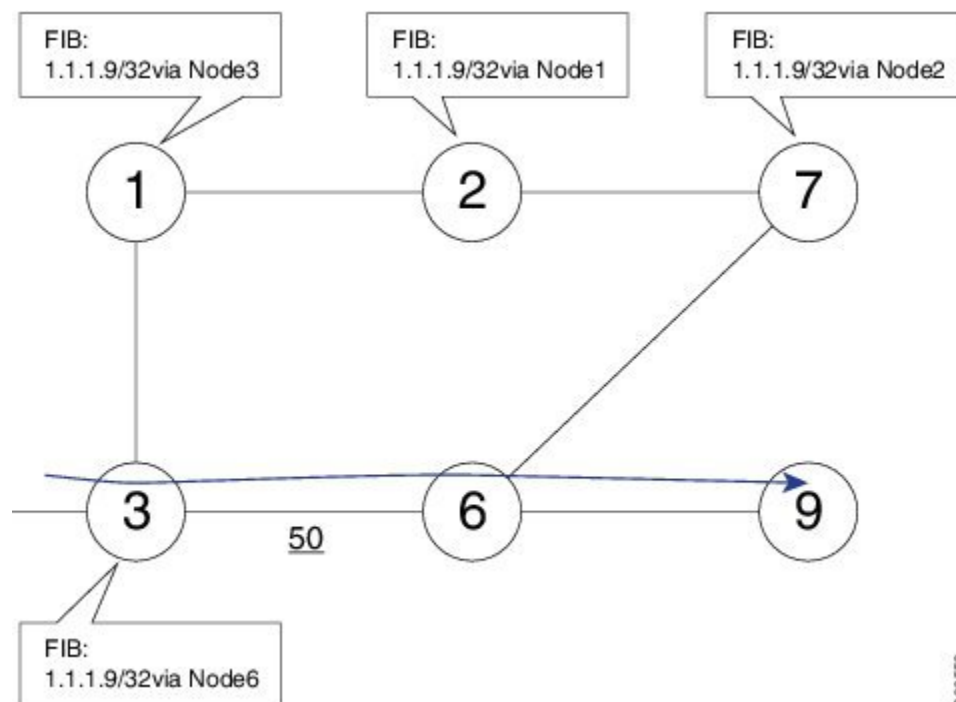
If a link is established between Node 6 and Node 7, the shortest path for a packet from Node 3 to Node 9 would be Node 1, Node 2, Node 7, and Node 6 before the packet reaches the destination, Node 9.

Figure 11: Microloop Example Topology—Shortest Path



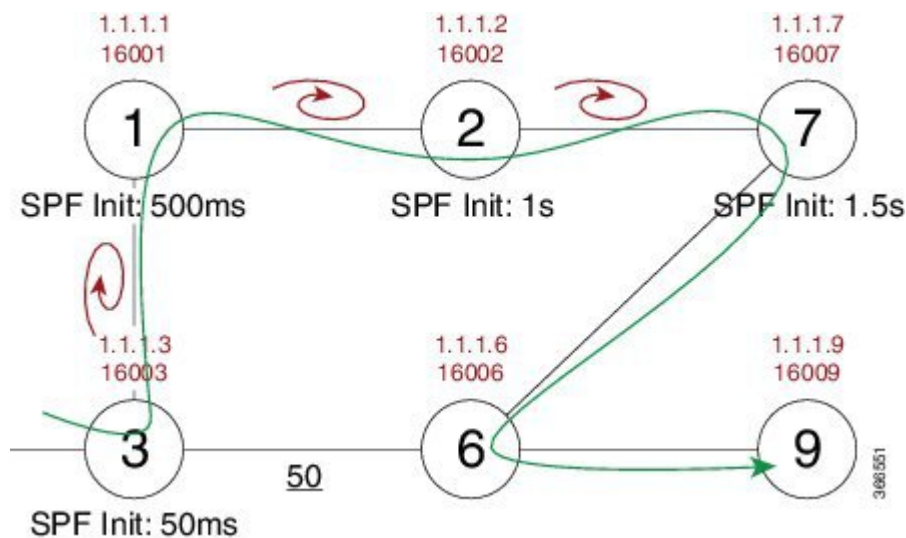
The following figure shows the Forwarding Information Base (FIB) table in each node before the link between Node 6 and Node 7 is established. The FIB entry contains the prefix of the destination node (Node 9) and the next hop.

Figure 12: Microloop Example Topology—FIB Entry



When the link between Node 6 and Node 7 comes up, microloops occur for the links based on the order of convergence of each node. In this example, Node 3 converges first with Node 1 resulting in a microloop between Node 3 and Node 1. Then, Node 1 converges next resulting in a microloop between Node 1 and Node 2. Next, Node 2 converges next resulting in a microloop between Node 2 and Node 7. Finally, Node 7 converges resolving the microloop and the packet reaches the destination Node 9, as shown in the following figure.

Figure 13: Microloop Example Topology—Microloops

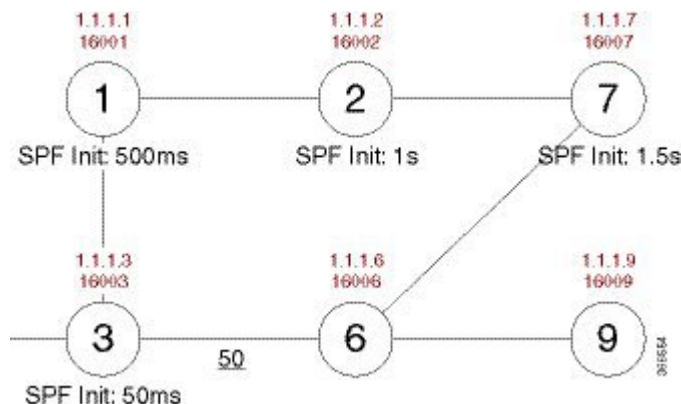


Adding the SPF convergence delay, microloop results in a loss of connectivity for 1.5 seconds, which is the convergence duration specified for node 7.

## How Segment Routing Prevents Microloops?

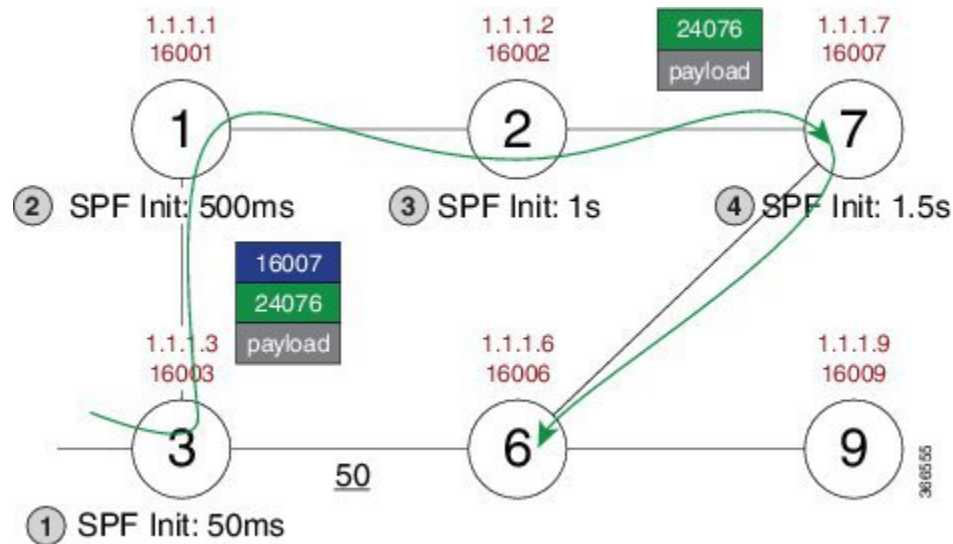
Using the example used to explain microloops, this section explains how to segment routing prevents microloops. Node 3 in the example is enabled with the **microloop avoidance segment-routing** command. After the link between Node 6 and Node 7 comes up, Node 3 computes a new microloop on the network.

Figure 14: Microloop Example Topology—Segment Routing



Instead of updating the FIB table, Node 3 builds a dynamic loop-free alternate (LFA) SR path for the destination (Node 9) using a list of segments IDs, which include the prefix segment ID (SID) of Node 7, which is 16007, and the adjacency segment ID (SID) of Node 6, which is 24076.





So, the SR path enables a packet from Node 3 reaches its destination Node 9, without the risk of microloop until the network converges. Finally, Node 3 updates the FIB for the new path.

Use the protected keyword with the **microloop avoidance segment-routing** command, to enable microloop avoidance for protected prefixes only. The **microloop avoidance rib-update-delay milliseconds** command can be used to configure the delay in milliseconds for a node to wait before updating the node's forwarding table and stop using the microloop avoidance policy. The default value for the RIB delay is 5000 milliseconds.

## How to Enable Segment Routing uLoop Avoidance

### Enabling ISIS - Microloop Avoidance

The following is a sample configuration code snippet to enable microloop avoidance in ISIS.

```
router isis 1
 net 49.0001.0000.0000.0209.00
 metric-style wide
 nsf cisco
 distribute link-state
 segment-routing mpls
 segment-routing prefix-sid-map advertise-local
 fast-reroute per-prefix level-1 all
 fast-reroute per-prefix level-2 all
 fast-reroute ti-lfa level-2
 microloop avoidance segment-routing
 bfd all-interfaces
 mpls traffic-eng router-id Loopback0
 mpls traffic-eng level-2
!
```

### Verifying ISIS - Microloop Avoidance

Use the **show isis private** command to check if the repair path exists or not.

```

Router# show isis private
Process: pdb 0x84BD34E0, name isis, tag 1, index -1, mask 0, lvl 3, handle 84BD34E0
Mode: ipv4 active-ip
idbqueue (4): Te0/0/27 Te0/0/26 Lo1 Lo0
passive_idbq (0):
IPv4:
  ip_pdb 0x84BFFE18, index 3, mask 0x8
  ipv6_idbq (0):
TID 0: topoid 0x00000000 (IPv4)
  num_L2 1, redist_L2 0
  ip_ipdb 0x84C00030
  ip_idbq (4): Lo0 Lo1 Te0/0/26 Te0/0/27

```

#### ISIS Microloop Avoidance Info:

enable state: enabled for segment-routing microloop avoidance

user config: microloop avoidance segment-routing

delay: 5000,

primary\_update\_complete\_pending: 0

```

L1:
  local_down_event[1]:
  uloop_event_if[1]:
  uloop_spf_on[1]:0, uloop_delay_reged[1]:0,
  uloop_rt_track_requested[1]:0
  uloop_delay_timer[1] is not running
  uloop_lsp_nbr[1]:
  Saved local neighbor list:

L2:
  local_down_event[2]:
  uloop_event_if[2]:
  uloop_spf_on[2]:0, uloop_delay_reged[2]:0,
  uloop_rt_track_requested[2]:0
  uloop_delay_timer[2] is not running
  uloop_lsp_nbr[2]:
  Saved local neighbor list:
  ISIS Microloop avoidance L(1) lost adjacency queue is empty
  ISIS Microloop avoidance L(2) lost adjacency queue is empty

```

## Enabling OSPF - Microloop Avoidance

The following is a sample configuration code snippet to enable microloop avoidance in ISIS.

```

router ospf 1
nsr
nsf ietf
segment-routing mpls
segment-routing prefix-sid-map advertise-local
fast-reroute per-prefix enable prefix-priority low
fast-reroute per-prefix ti-lfa
fast-reroute per-prefix ti-lfa area 0
microloop avoidance segment-routing
network 170.50.0.0 0.0.255.255 area 0
network 203.203.203.203 0.0.0.0 area 0
bfd all-interfaces
mpls traffic-eng router-id Loopback1
mpls traffic-eng area 0
!

```

## Verifying OSPF - Microloop Avoidance

Use the **show ip ospf segment-routing microloop-avoidance** command to check if the repair path exists or not.

```
Router#show ip ospf segment-routing microloop-avoidance
```

```
OSPF Router with ID (213.213.213.213) (Process ID 1)
```

```
SR Microloop Avoidance is configured, delay 5000 msec
```

```
Area with ID (0)
```

```
Base Topology (MTID 0)
```

```
SR Microloop Avoidance is enabled and not running
```

```
Last topology change details:
```

```
Near end: 213.213.213.213
```

```
Far end: 215.215.215.215
```

```
Event: Link Up
```





# CHAPTER 13

## EVPN-VPWS Single Homing on Segment Routing

EVPN-VPWS single homing is a BGP control plane solution for point-to-point services. It has the ability to forward traffic from one network to another using Ethernet Segment without MAC lookup.

EVPN VPWS single homing works on both IP and SR core. IP core is used to support BGP while the SR core is used to switch packets between the endpoints.

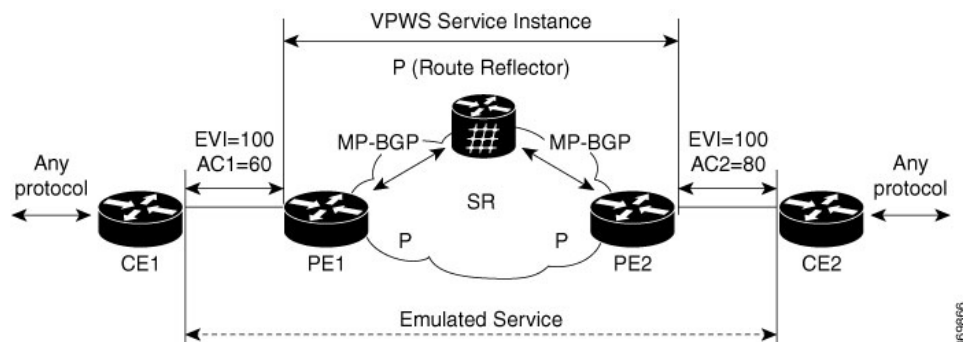
- [Information About EVPN-VPWS, on page 135](#)
- [Prerequisites for EVPN-VPWS, on page 136](#)
- [Restrictions for EVPN-VPWS, on page 136](#)
- [How to Configure EVPN-VPWS, on page 137](#)
- [Configuration Examples for EVPN-VPWS over SR, on page 143](#)
- [Additional References for EVPN-VPWS, on page 147](#)

## Information About EVPN-VPWS

The EVPN-VPWS solution supports per EVI Ethernet Auto Discovery route. EVPN defines a new BGP Network Layer Reachability Information (NLRI) that is used to carry all EVPN routes. BGP Capabilities Advertisement is used to ensure that two speakers support EVPN NLRI (AFI 25, SAFI 70) as per RFC 4760.

The architecture for EVPN VPWS is that the PEs run Multi-Protocol BGP in control-plane. The following image describes the EVPN-VPWS over SR configuration:

**Figure 15: EVPN-VPWS over SR Configuration**



## Benefits of EVPN-VPWS Single Homed

- Scalability is achieved without signaling pseudowires.
- There is ease of provisioning.
- Pseudowires (PWs) are not used.
- EVPN-VPWS Single Homed leverages BGP best-path selection (optimal forwarding).

## Prerequisites for EVPN-VPWS

- BGP must be configured for EVPN SAFI.
- Segment routing must be configured and made preferred.
- CE-facing interface such as service instance, is Ethernet family without IP address on PE.
- BGP session between PEs with 'address-family l2vpn evpn' to exchange EVPN routes.
- A BGP route reflector is supported and configured.
- An IGP such as, the ISIS is used for IP reachability between the PEs, whereas BGP is used for next-hop reachability

## Restrictions for EVPN-VPWS

- The combination of EVPN ID and VPWS Instance ID must be unique according to ASN.
- SR TE core is *not* supported.
- InterAS Option B is *not* supported.
- NSR is *not* supported for L2VPN family.
- Ensure NSF is configured on BGP, OSPF (iBGP), and MPLS.
- NSF is supported; however, neighbor flaps occur.
- Without NSF, when a Stateful Switchover (SSO) takes place, traffic drops for the L2VPN-EVP packets.
- **evpn vc stats** do *not* work in the **show l2vpn evpn vc id detail** command.
- ELB is *not* supported on EVPN.

### Scaling Information

Number of EVPN VPWS service instances supported is 4000.

# How to Configure EVPN-VPWS

The following steps are performed to configure EVPN-VPWS

- Configuring BGP for EVPN-VPWS
- Configuring EVPN-VPWS Instance

## Configuring BGP for EVPN-VPWS

To configure EVPN-VPWS in BGP, follow these steps:

### Procedure

```
router bgp 1
address-family l2vpn evpn
  neighbor 4.4.4.4 activate
  neighbor 4.4.4.4 send-community both
exit-address-family
```

## Configuring EVPN-VPWS Instance

To configure EVPN VPWS instance, follow these steps:

### Procedure

```
enable
configure terminal
l2vpn evpn instance 1 point-to-point
vpws context 1
  service target 1 source 1
  member GigabitEthernet0/0/6 service-instance 1
!
```

## Rewrite for EVI Service Instance

You must run the rewrite command when the VLANs are mismatched on the remote ACs. This allows ingress traffic movement. To configure EVPN-VPWS service instance for rewrite, follow these steps:

### Procedure

Rewrite on PE1:

```
interface GigabitEthernet0/0/6
service instance 1 ethernet
  encapsulation dot1q 1
!
```

Rewrite on CE1:

```
interface GigabitEthernet0/0/8
service instance 1 ethernet
  encapsulation dot1q 1
  rewrite ingress tag pop 1 symmetric
```

```
bridge-domain 1
```

# Configuring EVPN-VPWS for Logging

To configure EVPN-VPWS for logging, follow these steps:

## Procedure

```
enable
configure terminal
l2vpn evpn logging vc-state
end
```

# Verifying EVPN-VPWS Instance

## Verifying EVPN-VPWS Configuration

You can verify the configuration using the following show commands:

- **show l2vpn evpn vpws summary**
- **show l2vpn evpn vpws evi (<evpn-id> | all) [detail]**
- **show l2vpn evpn rib ead [per] |evi**
- **show ip bgp l2vpn evpn evi <evpn-id> context**
- **show l2vpn evpn memory**

This command displays a summary of L2VPN EVPN with total number of EVIs, VCs and routes.

```
show l2vpn evpn vpws summary
L2VPN EVPN VPWS:
  EVIs (point-to-point): 3
  Total VCs: 3
    3 up, 0 down, 0 admin-down, 0 hot-standby, 0 other
  BGP: ASN 1, address-family l2vpn evpn configured
  Router ID: 2.2.2.2
```

This command displays brief or detail info for EVIs.

```
show l2vpn evpn vpws evi 1
EVI      VCs      Up      Down      AdmDn      HotSby      Other
-----
1         1         1         0         0           0           0
PE-1#show l2vpn evpn vpws evi all
PE-1#show l2vpn evpn vpws evi all
```

```
EVI      VCs      Up      Down      AdmDn      HotSby      Other
-----
1         1         1         0         0           0           0
2         1         1         0         0           0           0
3         1         1         0         0           0           0
```

```
show l2vpn evpn vpws evi all detail
```

```
EVPN instance: 1 (point-to-point)
RD: 2.2.2.2:1 (auto)
```



```

Import-RTs: 1:1
Export-RTs: 1:1
Total VCs: 1
    1 up, 0 down, 0 admin-down, 0 hot-standby, 0 other

EVPN instance: 2 (point-to-point)
RD: 2.2.2.2:2 (auto)
Import-RTs: 1:2 100:100
Export-RTs: 1:2 100:100
Total VCs: 1
    1 up, 0 down, 0 admin-down, 0 hot-standby, 0 other

EVPN instance: 3 (point-to-point)
RD: 2.2.2.2:3 (auto)
Import-RTs: 1:3
Export-RTs: 1:3
Total VCs: 1
    1 up, 0 down, 0 admin-down, 0 hot-standby, 0 other

show l2vpn evpn vpws evi 1 detail

EVPN instance: 1 (point-to-point)
RD: 2.2.2.2:1 (auto)
Import-RTs: 1:1
Export-RTs: 1:1
Total VCs: 1
    1 up, 0 down, 0 admin-down, 0 hot-standby, 0 other

```

This command displays the contents of the global EVPN route.

```

show l2route evpn ead per-evi

```

EVI	ETag	Prod	ESI	Next Hop(s)	Label
1	1	L2VPN	N/A	N/A	19
1	1	BGP	0000.0000.0000.0000.0000	4.4.4.4	16
2	2	L2VPN	N/A	N/A	22
2	2	BGP	0000.0000.0000.0000.0000	4.4.4.4	24
3	3	L2VPN	N/A	N/A	25
3	3	BGP	0000.0000.0000.0000.0000	4.4.4.4	29

This command displays the contents of the global route-target (RT).

```

show ip bgp l2vpn evpn evi 1 context
EVI evi_1 (EVI Id = 1); default RD 2.2.2.2:1
Export EVPN route-target communities
RT:1:1
Import EVPN route-target communities
RT:1:1

```

This command displays brief or detail EVPN memory usage.

```

show l2vpn evpn memory

```

Allocator-Name	In-use/Allocated	Count
EVPN DB	900/32820 ( 2%)	[ 25] Chunk
EVPN MGR DB	108/32820 ( 0%)	[ 3] Chunk
EVPN Mgr EVI chunk	792/10052 ( 7%)	[ 3] Chunk
EVPN Mgr Msg chunk	0/8484 ( 0%)	[ 0] Chunk
EVPN Mgr Thread	4468/5040 ( 88%)	[ 11]
EVPN VPWS Thread	6304/6980 ( 90%)	[ 13]
EVPN context chunk	2268/32820 ( 6%)	[ 3] Chunk
EVPN context handle table	35480/35584 ( 99%)	[ 2]

```

EVPN dtrace elem per-cont :      10512/32820      ( 32%) [    292] Chunk
EVPN dtrace stridx       :      1194876/1194928    ( 99%) [     1]
EVPN dtrace stridx freeli :      132764/132816    ( 99%) [     1]
EVPN dtrace stridx hash  :           40/92        ( 43%) [     1]
EVPN dtrace stridx slots :      132764/132816    ( 99%) [     1]
EVPN dtrace stridx2slot  :      132764/132816    ( 99%) [     1]
EVPN instance chunk      :           504/8484      ( 5%) [     3] Chunk
EVPN rt-db ee            :           168/376      ( 44%) [     4]
EVPN rt-db rte           :           240/448      ( 53%) [     4]

```

Total allocated: 1.716 Mb, 1758 Kb, 1800196 bytes

## Verifying EVPN-VPWS Configuration for Logging

You can verify the logging using the **show l2vpn evpn vc** command.

This command displays brief information for VCs.

```

show l2vpn evpn vpws vc all

```

EVPN ID	Source	Target	Type	Name/Interface	Status
1	1	1	p2p	1	up
				Gi0/0/6:1	up
2	2	2	p2p	2	up
				Gi0/0/6:2	up
3	3	3	p2p	3	up
				Gi0/0/6:3	up

This command displays detail information for VCs.

```

show l2vpn evpn vpws vc all detail
EVPN name: 1, state: up, type: point-to-point
EVPN ID: 1
VPWS Service Instance ID: Source 1, Target 1
Labels: Local 19, Remote 16
Next Hop Address: 4.4.4.4
Associated member interface Gi0/0/6 up, Gi0/0/6:1 status is up
Dataplane:
  SSM segment/switch IDs: 4149/4139 (used), PWID: 3
Rx Counters
  27840 input transit packets, 3397212 bytes
  0 drops
Tx Counters
  27885 output transit packets, 3400188 bytes
  0 drops
17 VC FSM state transitions, Last 10 shown
RemUp: RemWait -> Act, Sun Sep 01 08:27:29.093 (2d07h ago)
DpUp: Act -> Est, Sun Sep 01 08:27:29.113 (2d07h ago)
LocDn: Est -> LocWait, Sun Sep 01 08:31:07.953 (2d07h ago)
LocUp: LocWait -> RemWait, Sun Sep 01 08:31:19.665 (2d07h ago)
RemUp: RemWait -> Act, Sun Sep 01 08:31:19.665 (2d07h ago)
DpUp: Act -> Est, Sun Sep 01 08:31:19.681 (2d07h ago)
LocDn: Est -> LocWait, Sun Sep 01 08:42:17.685 (2d07h ago)
LocUp: LocWait -> RemWait, Sun Sep 01 08:44:23.869 (2d07h ago)
RemUp: RemWait -> Act, Sun Sep 01 08:44:23.869 (2d07h ago)
DpUp: Act -> Est, Sun Sep 01 08:44:23.933 (2d07h ago)

EVPN name: 2, state: up, type: point-to-point
EVPN ID: 2
VPWS Service Instance ID: Source 2, Target 2
Labels: Local 22, Remote 24
Next Hop Address: 4.4.4.4
Associated member interface Gi0/0/6 up, Gi0/0/6:2 status is up
Dataplane:

```

```

SSM segment/switch IDs: 12342/8238 (used), PWID: 4
Rx Counters
  21214 input transit packets, 2503252 bytes
  0 drops
Tx Counters
  0 output transit packets, 0 bytes
  0 drops
17 VC FSM state transitions, Last 10 shown
  RemUp: RemWait -> Act, Sun Sep 01 08:27:29.093 (2d07h ago)
  DpUp: Act -> Est, Sun Sep 01 08:27:29.113 (2d07h ago)
  LocDn: Est -> LocWait, Sun Sep 01 08:31:07.961 (2d07h ago)
  LocUp: LocWait -> RemWait, Sun Sep 01 08:31:19.665 (2d07h ago)
  RemUp: RemWait -> Act, Sun Sep 01 08:31:19.665 (2d07h ago)
  DpUp: Act -> Est, Sun Sep 01 08:31:19.681 (2d07h ago)
  LocDn: Est -> LocWait, Sun Sep 01 08:42:17.689 (2d07h ago)
  LocUp: LocWait -> RemWait, Sun Sep 01 08:44:23.873 (2d07h ago)
  RemUp: RemWait -> Act, Sun Sep 01 08:44:23.873 (2d07h ago)
  DpUp: Act -> Est, Sun Sep 01 08:44:23.937 (2d07h ago)

EVPN name: 3, state: up, type: point-to-point
EVPN ID: 3
VPWS Service Instance ID: Source 3, Target 3
Labels: Local 25, Remote 29
Next Hop Address: 4.4.4.4
Associated member interface Gi0/0/6 up, Gi0/0/6:3 status is up
Dataplane:
  SSM segment/switch IDs: 20539/12344 (used), PWID: 5
Rx Counters
  2585 input transit packets, 305078 bytes
  0 drops
Tx Counters
  2591 output transit packets, 305258 bytes
  0 drops
8 VC FSM state transitions, Last 8 shown
  Prov: Idle -> Prov, Tue Sep 03 09:43:12.266 (06:39:15 ago)
  EviUp: Prov -> LocWait, Tue Sep 03 09:43:12.266 (06:39:15 ago)
  LocUp: LocWait -> RemWait, Tue Sep 03 09:43:12.266 (06:39:15 ago)
  RemUp: RemWait -> Act, Tue Sep 03 09:43:12.278 (06:39:15 ago)
  DpUp: Act -> Est, Tue Sep 03 09:43:12.294 (06:39:15 ago)
  RemDn: Est -> RemWait, Tue Sep 03 09:43:33.210 (06:38:54 ago)
  RemUp: RemWait -> Act, Tue Sep 03 09:43:33.234 (06:38:54 ago)
  DpUp: Act -> Est, Tue Sep 03 09:43:33.262 (06:38:54 ago)

```

## Troubleshooting

### Virtual Circuit (VC) is in Down state

EVPN VPWS protocol has no communication of VC state between endpoints. VC can be up on one end and down on the other end in the following cases:

- When the core-facing MPLS dataplane is down on one side only. For example, if loopback is configured with /24 on one-end but configured correctly with /32 at the other end.
- UUT has no remote EVPN EAD route from peer. This can happen in the following scenarios:
  - Peer never sent it.
  - Peer sent it, but RT mismatches. No intersection between UUT Import-RT and peer Export-RT.
  - Peer sent it, RT matches, but E-tag mismatches: For service E-tags, TGT/SRC, UUT has X/Y and peer has Y/Z.

**Problem** VC is in down state.

**Possible Cause** None

**Solution** Perform these steps to check whether the VC is not active:

- **Solution** Check if any VC is not active.
- **Solution** Identify EVIs that has not got an active VCs
- **Solution** Gather information for the EVIs that has not got an active VCs
- **Solution** Locate the inactive VCs for the EVI
- **Solution** Display detail information of the not active VC

### Solution

```
show l2vpn evpn vpws vc all detail
EVPN name: vc100, state: up, type: point-to-point
  EVPN ID: 100
  VPWS Service Instance ID: Source 1, Target 2
  Labels: Local 16, Remote 16

// Must have a valid Local Label. If missing, contact support.

// Must have valid Remote Label. If missing, then there is no matching remote route.
Cross-check with BGP: 'show bgp l2vpn evpn [...] detail'.

Next Hop Address: 192.168.0.3

// Must have valid Next Hop Address. If missing, then there is no matching remote route.
Cross-check with BGP: 'show bgp l2vpn evpn [...] detail'.

Associated member Et0/0 is up, status is up

// AC must be up. If not up, check why.

Dataplane:
  SSM segment/switch IDs: 4098/4097 (used), PWID: 1
Rx Counters
  78 input transit packets, 26425 bytes
  0 drops
Tx Counters
  79 output transit packets, 28240 bytes
  0 drops
5 VC FSM state transitions, Last 5 shown
  Prov: Idle -> Prov, Tue Sep 29 13:15:37.848 (00:52:21 ago)
  AdmUp: Prov -> LocWait, Tue Sep 29 13:15:40.287 (00:52:18 ago)
  LocUp: LocWait -> RemWait, Tue Sep 29 13:15:40.287 (00:52:18 ago)
  RemUp: RemWait -> Act, Tue Sep 29 13:17:19.368 (00:50:39 ago)
  DpUp: Act -> Est, Tue Sep 29 13:17:19.371 (00:50:39 ago)

// Pay close attention to last line of VC FSM history. The format is:
// <Event>: <OldState> -> <NewState>
// Troubleshooting info appears below.
```

## VC FSM History

**Problem** The state of the VC is Prov — Provisioned: VC is disabled.

**Possible Cause** None

**Solution** Perform these steps for a solution to the state:

- **Solution** Check BGP is running.
- **Solution** Check BGP 'address-family l2vpn evpn' is configured.
- **Solution** Check VC is not shutdown.

**Problem** The state of the VC is LocWait — Local-Wait: Waiting for local AC information to come up.

**Possible Cause** None

**Solution** Check AC is up.

**Problem** The state of the VC is Act — Activating: Control plane ok. Trying to activate dataplane.

**Possible Cause** None

- **Solution** Check core facing information is up.
- **Solution** Check Segment-Routing is configured and preferred.

## Remote-Wait State

**Problem** The state of the VC is RemWait — Remote-Wait: Waiting for matching remote route.

**Possible Cause** This state occurs due to no matching remote route for the VC. A matching remote route means all of the following are true:

- Route is present in BGP. Requires a local EVI to have route target in the route.
- Remote path is best path.
- Route is present in global EVPN route.
- Route is present in EVI route. Requires the EVI to have route target in the route.
- Route has ETag which matches the VC source identity. (**service target <tgt-id> source <src-id>**).

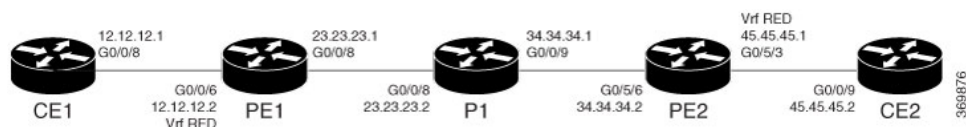
**Solution** Perform these steps to check whether the VC is in remote wait state:

### Solution

- Check for EVI configuration mismatch.
- Check for VC configuration mismatch.
- Check if the remote route is present in BGP.
  - If no remote route then check if
    - remote route was discarded by BGP due to RT filter
    - peer did not send route to UUT
    - EVI or VC configuration mismatch
    - all the prerequisites are satisfied
  - If a remote route is present in global EVPN then check if the remote route is present in EVI route.
    - **Solution** Check for EVI or VC configuration mismatch.

## Configuration Examples for EVPN-VPWS over SR

The following example has running configurations on PE1 and PE2:



### Example: EVPN-VPWS over SR, PE-1 configuration

```
show run int lo0
Building configuration...
```

```
Current configuration : 63 bytes
!
interface Loopback0
ip address 2.2.2.2 255.255.255.255
end
```

```
PE-1#show run | sec router ospf 1
router ospf 1
router-id 2.2.2.2
segment-routing mpls
segment-routing prefix-sid-map advertise-local
network 2.2.2.2 0.0.0.0 area 0
network 23.23.23.0 0.0.0.255 area 0
bfd all-interfaces
PE-1#
```

```
segment-routing mpls
!
set-attributes
  address-family ipv4
  sr-label-preferred
  exit-address-family
!
!
connected-prefix-sid-map
  address-family ipv4
  2.2.2.2/32 index 1 range 1
  exit-address-family
!
```

```
l2vpn evpn instance 1 point-to-point
vpws context 1
  service target 1 source 1
  member GigabitEthernet0/0/6 service-instance 1
!
l2vpn evpn instance 2 point-to-point
route-target export 100:100
route-target import 100:100
vpws context 2
  service target 2 source 2
  member GigabitEthernet0/0/6 service-instance 2
!
l2vpn evpn instance 3 point-to-point
vpws context 3
  service target 3 source 3
  member GigabitEthernet0/0/6 service-instance 3
!
```

```
PE-1#show run | sec router bgp 1
router bgp 1
bgp log-neighbor-changes
neighbor 4.4.4.4 remote-as 1
```

```

neighbor 4.4.4.4 update-source Loopback0
!
address-family vpnv4
  neighbor 4.4.4.4 activate
  neighbor 4.4.4.4 send-community both
exit-address-family
!
address-family l2vpn evpn
  neighbor 4.4.4.4 activate
  neighbor 4.4.4.4 send-community both
exit-address-family
!
address-family ipv4 vrf RED
  redistribute ospf 2
exit-address-family
PE-1#
interface GigabitEthernet0/0/6
#CE facing interface
no ip address
media-type sfp
negotiation auto
no keepalive
service instance 1 ethernet
  encapsulation dot1q 1
!
service instance 2 ethernet
  encapsulation dot1q 2
!
service instance 3 ethernet
  encapsulation untagged
!
service instance 12 ethernet
  encapsulation dot1q 12
  rewrite ingress tag pop 1 symmetric
  bridge-domain 12
!
end

```

```

PE-1#show run int gigabitEthernet 0/0/8
Building configuration...

```

```

Current configuration : 156 bytes
!
interface GigabitEthernet0/0/8
#core facing interface
ip address 23.23.23.1 255.255.255.0
ip ospf 1 area 0
media-type sfp
negotiation auto
mpls ip
bfd template templ
end

```

### Example: EVPN-VPWS over SR, PE2 configurations:

```

show run int lo0
Building configuration...

```

```

Current configuration : 63 bytes
!
interface Loopback0
ip address 4.4.4.4 255.255.255.255

```

```

end

PE-2#show run | sec router ospf 1
router ospf 1
router-id 4.4.4.4
segment-routing mpls
segment-routing prefix-sid-map advertise-local
network 4.4.4.4 0.0.0.0 area 0
network 34.34.34.0 0.0.0.255 area 0
bfd all-interfaces
PE-2#show run | sec segment-r
segment-routing mpls
!
set-attributes
  address-family ipv4
    sr-label-preferred
  exit-address-family
!
!
connected-prefix-sid-map
  address-family ipv4
    4.4.4.4/32 index 1 range 1
  exit-address-family
!
PE-2#show run | sec l2vpn evpn
address-family l2vpn evpn
  neighbor 2.2.2.2 activate
  neighbor 2.2.2.2 send-community both
l2vpn evpn instance 1 point-to-point
vpws context 1
  service target 1 source 1
  member GigabitEthernet0/5/3 service-instance 1
!
l2vpn evpn instance 2 point-to-point
vpws context 2
  service target 2 source 2
  member GigabitEthernet0/5/3 service-instance 2
!
l2vpn evpn instance 3 point-to-point
vpws context 3
  service target 3 source 3
  member GigabitEthernet0/5/3 service-instance 3
!
PE-2# show run | sec router bgp 1
router bgp 1
bgp log-neighbor-changes
neighbor 2.2.2.2 remote-as 1
neighbor 2.2.2.2 update-source Loopback0
!
address-family ipv4
  neighbor 2.2.2.2 activate
  neighbor 2.2.2.2 send-community both
exit-address-family
!
address-family vpnv4
  neighbor 2.2.2.2 activate
  neighbor 2.2.2.2 send-community both
exit-address-family
!
address-family l2vpn evpn
  neighbor 2.2.2.2 activate
  neighbor 2.2.2.2 send-community both
exit-address-family
!

```



```
address-family ipv4 vrf RED
  redistribute ospf 2
exit-address-family
PE-2#

PE-2#show run int gigabitEthernet 0/5/3
Building configuration...

Current configuration : 369 bytes
!
interface GigabitEthernet0/5/3
#access facing interface
no ip address
negotiation auto
no keepalive
service instance 1 ethernet
  encapsulation dot1q 1
!
service instance 2 ethernet
  encapsulation dot1q 2
!
service instance 3 ethernet
  encapsulation untagged
!
service instance 45 ethernet
  encapsulation dot1q 45
  rewrite ingress tag pop 1 symmetric
  bridge-domain 45
!
end

PE-2#show run int gigabitEthernet 0/5/6
Building configuration...

Current configuration : 140 bytes
!
interface GigabitEthernet0/5/6
#core facing interface
ip address 34.34.34.2 255.255.255.0
ip ospf 1 area 0
negotiation auto
mpls ip
bfd template templ
end
```

## Additional References for EVPN-VPWS

### Related Documents

Related Topic	Document Title
Cisco IOS commands	<a href="#">Cisco IOS Master Commands List, All Releases</a>

**Standards and RFCs**

Standard/RFC	Title
RFC 7432	<i>BGP MPLS-Based Ethernet VPN</i>
Standard	<i>VPWS support in EVPN</i>

**MIBs**

MIB	MIBs Link
•	—

**Technical Assistance**

Description	Link
<p>The Cisco Support website provides extensive online resources, including documentation and tools for troubleshooting and resolving technical issues with Cisco products and technologies.</p> <p>To receive security and technical information about your products, you can subscribe to various services, such as the Product Alert Tool (accessed from Field Notices), the Cisco Technical Services Newsletter, and Really Simple Syndication (RSS) Feeds.</p> <p>Access to most tools on the Cisco Support website requires a Cisco.com user ID and password.</p>	<a href="http://www.cisco.com/cisco/web/support/index.html">http://www.cisco.com/cisco/web/support/index.html</a>



## CHAPTER 14

# Configure Performance Measurement

**Table 3: Feature History**

Feature Name	Release Information	Description
Segment Routing Performance Measurement Delay Measurement Using RFC 5357 (TWAMP Light)	Cisco IOS XE Amsterdam 17.3.1	This feature enables hardware timestamping. The Performance Measurement (PM) for link delay uses the light version of Two-Way Active Measurement Protocol (TWAMP) over IP and UDP defined in Appendix I of RFC 5357. TWAMP provides an alternative for interoperability when RFC 6374 is not used.

Network performance metrics such as packet loss, delay, delay variation, and bandwidth utilization are a critical measure for traffic engineering (TE) in service provider networks. These metrics provide network operators with information about characteristics of their networks for performance evaluation and helps to ensure compliance with service level agreements. The service-level agreements (SLAs) of service providers depend on the ability to measure and monitor these network performance metrics. Network operators can use the performance measurement (PM) feature to monitor the network metrics for links as well as end-to-end TE label switched paths (LSPs).

The following table explains the functionalities supported by the performance measurement feature for measuring delay for links or SR policies.

**Table 4: Performance Measurement Functionalities**

Functionality	Details
Profiles	Configure different profiles for different types of delay measurements. Delay profile type interfaces are used for link-delay measurement. Delay profile type sr-policy is used for SR policy delay measurements. Delay profile allows you to schedule probe and configure metric advertisement parameters for delay measurement.

Functionality	Details
Probe and burst scheduling	Schedule probes and configure metric advertisement parameters for delay measurement.
Metric advertisements	Advertise measured metrics periodically using configured thresholds. Also supports accelerated advertisements using configured thresholds.
Measurement history and counters	Maintain packet delay and loss measurement history and also session counters and packet advertisement counters.

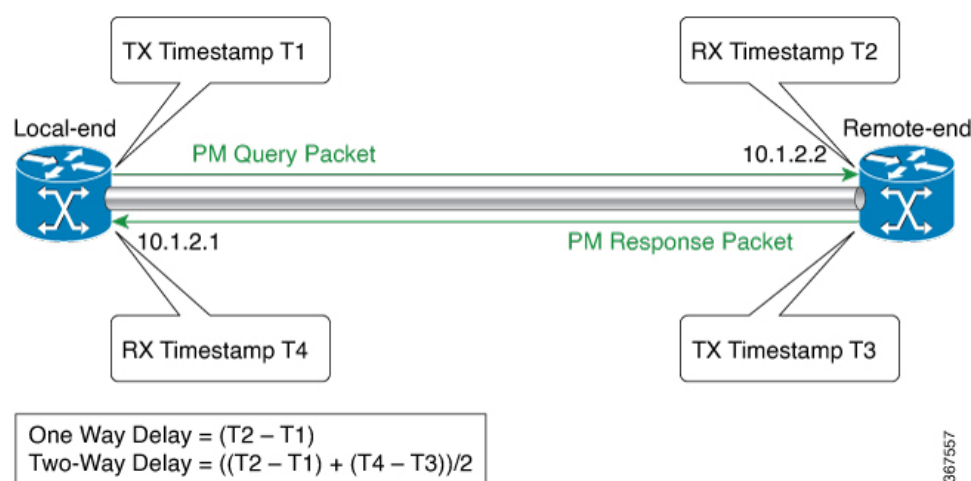
- [Link Delay Measurement](#), on page 150
- [End-to-End Delay Measurement](#), on page 154
- [SR-PCE: Enabling SR PM Delay or Liveness for PCE-Initiated Policies](#), on page 158
- [Telemetry \(Model-Based Telemetry and Event-Based Telemetry\) Support for Performance Measurement](#), on page 162
- [Configuring UDP Destination Port](#), on page 168

## Link Delay Measurement

The PM for link delay uses the light version of Two-Way Active Measurement Protocol (TWAMP) over IP and UDP defined in Appendix I of RFC 5357. Hence, only TWAMP test sessions are implemented and not the TWAMP control protocol. TWAMP provides an alternative for interoperability when RFC 6374 is not used. TWAMP packets are carried over IP and UDP. Thus, the dependency on MPLS dataplane is eliminated.

The following figure explains the PM query and response for link delay.

**Figure 16: Performance Measurement for Link Delay**



The PM query and response for link delay can be described in the following steps:

1. The local-end router sends PM query packets periodically to the remote side once the egress line card on the router applies timestamps on packets.
2. Ingress line card on the remote-end router applies time-stamps on packets as soon as they are received.

3. The remote-end router sends the PM packets containing time-stamps back to the local-end router. The remote-end router time-stamps the packet just before sending it for two-way measurement.
4. The local-end router time-stamps the packet as soon as the packet is received for two-way measurement.
5. One-way delay and optionally two-way delay is measured using the time-stamp values in the PM packet.

## Restrictions and Usage Guidelines for PM for Link Delay

The following restrictions and guidelines apply for the PM for link delay feature for different links.

- For broadcast links, only point-to-point (P2P) links are supported. P2P configuration on IGP is required for flooding the value.
- Only TWAMP protocol based PM probes are supported. MPLS-GAL based PM probes are not supported.
- For one-way delay measurement, clocks should be synchronized on two end-point nodes of the link using PTP.

## PM Link Delay: Default Values for Different Parameters

The default values for the different parameters in the PM for link delay is given as follows:

- probe: The default mode for probe is two-way delay measurement. If you are configuring one-way delay measurement, hardware clocks must be synchronized between the local-end and remote-end routers using precision time protocol (PTP).
- interval: The default probe interval is 30 seconds. The range is from 30 to 3600 seconds.
- burst count: The default value is 10 and range is from 1 to 30.
- burst interval: The default value is 3000 milliseconds and the range is from 30 to 15000 milliseconds.
- periodic advertisement: Periodic advertisement is enabled by default.
- periodic-advertisement interval: The default value is 120 seconds and the interval range is from 30 to 3600 seconds.
- periodic-advertisement threshold: The default value of periodic advertisement threshold is 10 percent.
- periodic-advertisement minimum: The default value is 1000 microseconds (usec) and the range is from 0 to 100000 microseconds.
- accelerated advertisement: Accelerated advertisement is disabled by default.
- accelerated-advertisement threshold: The default value is 20 percent and the range is from 0 to 100 percent.
- accelerated-advertisement minimum: The default value is 1000 microseconds and the range is from 1 to 100000 microseconds.

## Configuration Example: PM for Link Delay

This example shows how to configure performance-measurement functionalities for link delay as a global default profile.

```
R1(config)#performance-measurement
R1(config-perf-meas)#delay-profile interfaces
R1(config-pm-dm-intf)#probe
R1(config-pm-dm-intf-probe)#interval 40
R1(config-pm-dm-intf-probe)#protocol twamp-light
R1(config-pm-dm-intf-probe)#burst count 5
R1(config-pm-dm-intf-probe-burst)#interval 40
R1(config-pm-dm-intf-probe-burst)#exit
R1(config-pm-dm-intf-probe)#exit
R1(config-pm-dm-intf)#advertisement periodic
R1(config-pm-dm-intf-adv-per)#interval 100
R1(config-pm-dm-intf-adv-per)#threshold 80
R1(config-pm-dm-intf-adv-per)#minimum-change 5000
R1(config-pm-dm-intf-adv-per)#exit
R1(config-pm-dm-intf)#advertisement accelerated
R1(config-pm-dm-intf-adv-acc)#threshold 30
R1(config-pm-dm-intf-adv-acc)#minimum-change 1100
R1(config-pm-dm-intf-adv-acc)#exit
```

This example shows how to enable PM for link delay over an interface.

```
R1(config)#performance-measurement
R1(config-perf-meas)#interface gigabitEthernet0/3/3
R1(config-pm-intf)#delay-measurement
R1(config-pm-intf-dm)#next-hop ipv4 170.50.62.1
R1(config-pm-intf)#exit
```

## Verification: PM Link Delay Configuration

This example shows how to use the **show performance-measurement summary [detail]** command to verify the PM for link-delay configuration.

```
R1#show performance-measurement summary detail
Total interfaces                : 3
Maximum PPS                    : 100 pkts/sec

Interface Delay-Measurement:
  Total sessions                : 3
  Profile configuration:
    Measurement Type            : Two-Way
    Computation interval        : 30 seconds
    Burst interval              : 3000 mSec
    Burst count                 : 10 packets
    Protocol                    : TWAMP-Lite Unauth
    HW Timestamp Supported      : No
    Periodic advertisement      : Enabled
      Interval                  : 30 (effective: 30) sec
      Threshold                 : 100%
      Minimum-Change            : 100000 uSec
    Accelerated advertisement   : Enabled
      Threshold                 : 100%
      Minimum-Change            : 100000 uSec
      Threshold crossing check   : Minimum-delay
  Counters:
```

```

Packets:
  Total sent                : 293020
  Total received            : 293016
Errors:
  TX:
    Total interface down    : 0
    Total no MPLS caps      : 0
    Total no IP address      : 0
    Total other              : 19
  RX:
    Total negative delay     : 144
    Total delay threshold exceeded : 0
    Total missing TX timestamp : 0
    Total missing RX timestamp : 0
    Total probe full         : 0
    Total probe not started   : 0
    Total control code error  : 0
    Total control code notif  : 0
Probes:
  Total started              : 29306
  Total completed            : 29155
  Total incomplete           : 148
  Total advertisements       : 3

Global Delay Counters:
  Total packets sent         : 293020
  Total query packets received : 293016
  Total invalid session id    : 0
  Total no session           : 0

HW Support for MPLS-GAL [RFC6374] timestamp : No
HW Support for TWAMP [RF5357] timestamp     : No
HW Support for 64 bit timestamp              : No
HW Support for IPv4 UDP Cheksum              : No

```

This example shows how to use the **show performance-measurement interfaces** *[interface-name]* **[detail]** command to verify the PM for link-delay configuration.

```

R1#show performance-measurement interfaces detail
Interface Name: GigabitEthernet0/2/3 (ifh: 0xA)
  Delay-Measurement          : Enabled
  Local IPV4 Address         : 170.50.62.2
  Local IPV6 Address         : ::
  State                      : Up

Delay Measurement session:
  Session ID                 : 1

Last advertisement:
  Advertised at: 09:21:08 12 2019 (439879 seconds ago)
  Advertised reason: Advertise delay config
  Advertised delays (uSec): avg: 2000, min: 2000, max: 2000, variance: 0

Next advertisement:
  Check scheduled at the end of the current probe (roughly every 30 seconds)
  No probes completed
  Rolling average (uSec): 3146

Current Probe:
  Started at 11:32:17 17 2019 (10 seconds ago)
  Packets Sent: 4, received: 4
  Measured delays (uSec): avg: 1999, min: 1500, max: 2499, variance: 499
  Probe samples:
    Packet Rx Timestamp Measured Delay
    11:32:17 17 2019 1999999

```

```

11:32:20 17 2019 1500000
11:32:23 17 2019 2499999
11:32:26 17 2019 1999999
Next probe scheduled at 11:32:46 17 2019 (in 19 seconds)
Next burst packet will be sent in 1 seconds
R1#

```

You can also use the following commands for verifying the PM for link delay on the local-end router.

Command	Description
<b>show performance-measurement history interfaces</b> <i>[interface]</i> probe	Displays the PM link-delay probe history for interfaces.
<b>show performance-measurement history interfaces</b> <i>[interface]</i> aggr	Displays the PM link-delay aggregated history for interfaces.
<b>show performance-measurement counters</b> <i>[interface interface]</i>	Displays the PM link-delay session counters.

You can also use the following commands for verifying the PM for link-delay configuration on the remote-end router.

Command	Description
<b>show performance-measurement responder summary</b>	Displays the PM for link-delay summary on the remote-end router (responder).
<b>show performance-measurement responder interfaces</b> <i>[interface]</i>	Displays PM for link-delay for interfaces on the remote-end router.
<b>show performance-measurement responder counters</b> <i>[interface interface]</i>	Displays the PM link-delay session counters on the remote-end router.

## End-to-End Delay Measurement

*Table 5: Feature History*

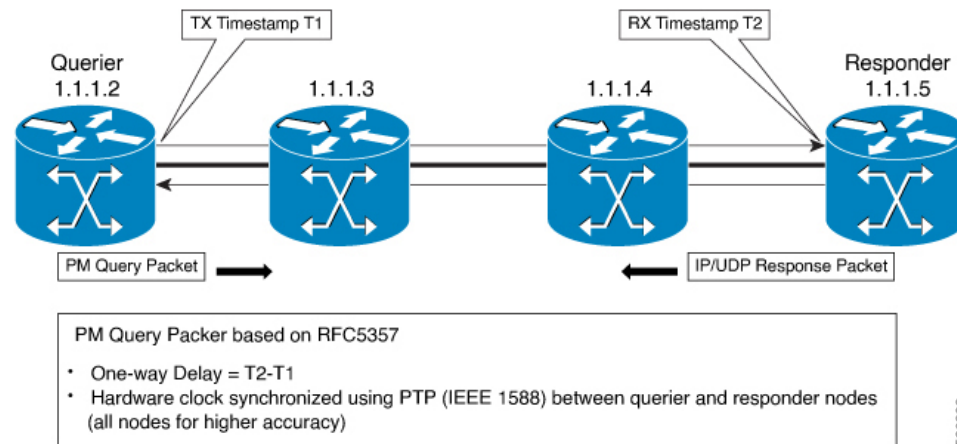
Feature Name	Release Information	Description
Segment Routing Performance Measurement End-to-End Delay Measurement	Cisco IOS XE Amsterdam 17.3.1	This feature allows to monitor the end-to-end delay experienced by the traffic sent over a Segment Routing policy. This feature ensures the delay does not exceed the specified threshold value and violate the SLAs. Use this feature to apply extended TE link delay metric (minimum delay value) to compute paths for Segment Routing policies as an optimization metric or as an accumulated delay bound.



Starting from Cisco IOS XE Release 17.3.1, end-to-end delay measurement feature is introduced for Segment Routing Performance Management. Use this feature to monitor the end-to-end delay experienced by the traffic sent over a Segment Routing policy. This feature ensures the delay does not exceed the specified threshold value and violate the SLAs. You can verify the end-to-end delay values before activating the candidate-path or the segment-list of the Segment Routing policy in the forwarding table. You can also use the end-to-end delay values to deactivate the active candidate-path or the segment-list of the Segment Routing Policy in the forwarding table. Use this feature to apply extended TE link delay metric (minimum delay value) to compute paths for Segment Routing policies as an optimization metric or as an accumulated delay bound.

The following figure explains the PM query and response for end-to-end delay measurement.

**Figure 17: Performance Measurement for End-to-End Delay Measurement**



The PM query and response for end-to-end delay measurement can be described in the following steps:

1. The querier router sends PM query packets periodically to the responder router once the egress line card on the router applies timestamps on packets.
2. Ingress line card on the responder router applies time-stamps on packets when they are received.
3. The end-to-end delay value of an SR Policy is different than the path computation result (the sum of TE link delay metrics) due to several factors like queuing delay within the routers.

## Configuration Example: PM for End-to-End Delay Management

These examples show how to configure on-demand segment routing policy for end-to-end delay management.

```
#show running-config | s on-demand color 800
on-demand color 800 -----> SR ODN Policy
authorize
performance-measurement -----> SR PM CLI
delay-measurement -----> SR PM CLI
candidate-paths
preference 1
constraints
segments
dataplane mpls
!
!
dynamic
pcep
metric
```

```

type delay
!
!
!
#

#show segment-routing traffic-eng policy name *216.216.216.216|800

Name: *216.216.216.216|800 (Color: 800 End-point: 216.216.216.216)
Owners : BGP
Status:
Admin: up, Operational: up for 01:27:24 (since 11-29 04:41:36.053)
Candidate-paths:
Preference 1 (BGP):
Dynamic (pce 12.12.12.12) (active)
Weight: 0, Metric Type: DELAY
Metric Type: DELAY, Path Accumulated Metric: 330
16011 [Prefix-SID, 205.205.205.205]
1133 [Adjacency-SID, 170.50.72.1 - 170.50.72.2]
16009 [Prefix-SID, 216.216.216.216]
Attributes:
Binding SID: 1218
Allocation mode: dynamic
State: Programmed
IPv6 caps enabled
#

```

This example shows how to configure performance-measurement functionalities for end-to-end delay management as a global default profile.

```

R1(config)#performance-measurement
R1(config-perf-meas)#delay-profile sr-policy
R1(config-pm-dm-intf)#probe
R1(config-pm-dm-intf-probe)#interval 40
R1(config-pm-dm-intf-probe)#protocol twamp-light
R1(config-pm-dm-intf-probe)#burst count 5
R1(config-pm-dm-intf-probe-burst)#interval 40
R1(config-pm-dm-intf-probe-burst)#exit
R1(config-pm-dm-intf-probe)#exit
R1(config-pm-dm-intf)#advertisement periodic
R1(config-pm-dm-intf-adv-per)#interval 100
R1(config-pm-dm-intf-adv-per)#threshold 80
R1(config-pm-dm-intf-adv-per)#minimum-change 5000
R1(config-pm-dm-intf-adv-per)#exit
R1(config-pm-dm-intf)#advertisement accelerated
R1(config-pm-dm-intf-adv-acc)#threshold 30
R1(config-pm-dm-intf-adv-acc)#minimum-change 1100
R1(config-pm-dm-intf-adv-acc)#exit

```

This example shows how to enable PM for end-to-end delay management over an interface.

```

R1(config)#performance-measurement
R1(config-perf-meas)#interface gigabitEthernet0/3/3
R1(config-pm-intf)#delay-measurement
R1(config-pm-intf-dm)#next-hop ipv4 170.50.62.1
R1(config-pm-intf)#exit

```

## Verification: PM End-to-End Delay Management Configuration

This example shows how to use the **show performance-measurement summary** command to verify the PM for end-to-end delay management configuration.

```

R1#show performance-measurement summary
Total interfaces                : 6
Total SR Policies               : 1
Maximum PPS                    : 1000 pkts/sec

SR Policy Delay-Measurement:
Total sessions                  : 1
Profile configuration:
  Measurement Type              : One-Way
  Computation interval          : 30 seconds
  Burst interval                : 3000 mSec
  Burst count                   : 10
  Protocol                     : TWAMP-Lite Unauth
  HW Timestamp Supported        : Yes
  Periodic advertisement       : Enabled
    Interval                    : 30 (effective: 30) sec
    Threshold                   : 15%
    Minimum-Change              : 600 uSec
  Accelerated advertisement     : Enabled
    Threshold                   : 25%
    Minimum-Change              : 900 uSec
  Threshold crossing check      : Minimum-delay
Counters:
  Packets:
    Total sent                  : 334
    Total received              : 0
  Errors:
    Total sent errors           : 0
    Total received errors       : 0
  Probes:
    Total started               : 33
    Total completed             : 0
    Total incomplete            : 33
    Total advertisements        : 0

Global Delay Counters:
Total packets sent              : 1251
Total query packets received    : 917
Total invalid session id        : 0
Total no session                : 0

HW Support for MPLS-GAL [RFC6374] timestamp : No
HW Support for TWAMP [RF5357] timestamp     : Yes
HW Support for 64 bit timestamp             : Yes
HW Support for IPv4 UDP Cheksum             : No
R1#

```

# SR-PCE: Enabling SR PM Delay or Liveness for PCE-Initiated Policies

*Table 6: Feature History*

Feature Name	Release Information	Description
SR-PCE: Enabling SR PM Delay or Liveness for PCE-Initiated Policies	Cisco IOS XE Bengaluru 17.6.1	This feature enables the Path Computation Element (PCE) that can provision a Segment Routing Traffic Engineering (SR-TE) policy to mitigate link congestion. Prior to this release, you could only enable PM link and delay measurement using CLI-based policies. Starting with this release, you can also use PCE to enable PM link and delay measurement.

The Path Computation Element (PCE) can provision a Segment Routing Traffic Engineering (SR-TE) policy to mitigate link congestion. The Path Computation Element Protocol (PCEP) describes a set of procedures by which a path computation client (PCC) can report and delegate control of head-end label switched paths (LSPs) sourced from the PCC to a PCE peer. The PCE can request the PCC to update and modify parameters of LSPs it controls. The stateful model also enables a PCC to allow the PCE to initiate computations allowing the PCE to perform network-wide orchestration. SR-PCE learns topology information by way of IGP (OSPF or IS-IS) or through BGP Link-State (BGP-LS).

Prior to Cisco IOS XE Bengaluru Release 17.6.1, you could only enable PM link and delay measurement using CLI-based policies. Starting with Cisco IOS XE Bengaluru Release 17.6.1, you can also use PCE to enable PM link and delay measurement.

Autoroute announcement is a steering mechanism in which IGP's automatically use the policy for destination's downstream of the policy end point. Autoroute announcement is performed using Cisco Crossworks Optimization Engine (COE). COE provides real-time network optimization allowing operators to maximize network utilization effectively and increase service velocity.

A PCE collects various pieces of network information to determine traffic flows causing link congestion. The PCE computes a suitable path to divert those flows and to alleviate the congestion. The PCE then deploys the SR-TE policy to divert the traffic leading to the congestion using the Stateful Path Computation Element Protocol (PCEP) to provision the policy. When the congestion is alleviated, the SR-TE policy is removed.

The PCEP message contains SID list to be deployed by the head-end. Path Computation Client (PCC) profiles allow activation of autoroute announce for the policy provisioned by PCEP, using the profile IDs. The profile ID on the PCE and PCC should match, otherwise the policy is not provisioned. For example, if the PCE provisions a policy with profile ID 1 and the head-end where the policy is being provisioned also has the PCC profile ID 1 configured with autoroute announce, COE-PCE initiated SR policy is activated for that policy.

SR-PCE is capable of computing paths using the following methods:

- TE metric—SR-PCE uses the TE metric in its path calculations to optimize cumulative TE metric.
- IGP metric—SR-PCE uses the IGP metric in its path calculations to optimize reachability.

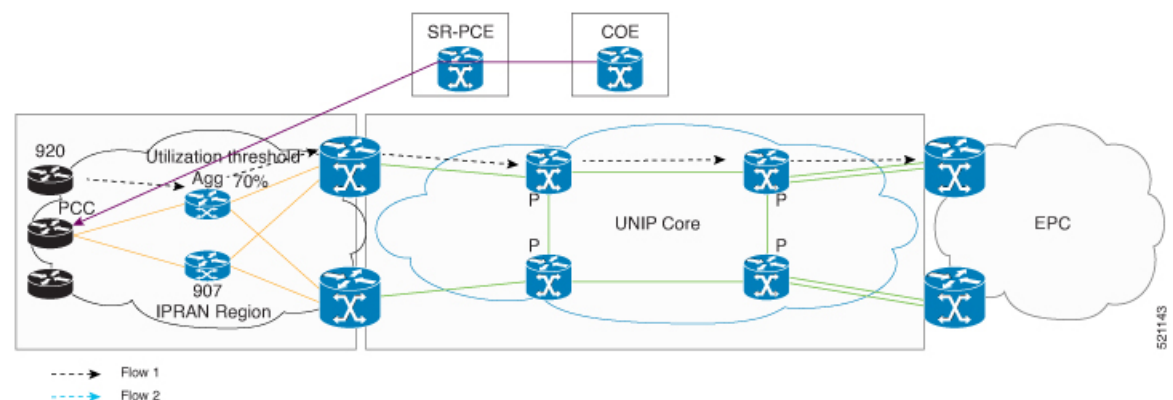
- **LSP Disjointness**—SR-PCE uses the path computation algorithms to compute a pair of disjoint LSPs. The disjoint paths can originate from the same head-end or different head-ends. Disjoint level refers to the type of resources that should not be shared by the two computed paths. SR-PCE supports the following disjoint path computations:
  - **Link** – Specifies that links are not shared on the computed paths.
  - **Node** – Specifies that nodes are not shared on the computed paths.
  - **SRLG** – Specifies that links with the same SRLG value are not shared on the computed paths.
  - **SRLG-node** – Specifies that SRLG and nodes are not shared on the computed paths.

## COE-PCE Initiated SR Policy

The following topology shows how an SR-PCE policy is initiated from COE:

- SR policy is configured on the COE with profile ID.
- COE pushes the SR policy to PCE and PCE forwards the SR policy to PCC.
- Profile ID on PCC is matched with the profile ID on COE-PCE.
- OSPF autoroute announce is configured on the PCC.
- The policy gets provisioned.
- The data traffic now adheres to the SR policy that is pushed from the COE.
- Complete SR Policy manipulation occurs only on COE.

**Figure 18: COE-PCE Initiated SR Policy**



## Configure SR-PCE: Enabling SR-PM Delay or Liveness for PCE-Initiated Policies

To enable SR-PM delay or liveness for PCE-Initiated policies, configure PCC and PCE nodes.

### Configure PCC Node:

To configure PCC node:

```

pcc
  pce address 9.9.9.9 source-address 1.1.1.1
  report-all
  profile 1
    autoroute
    include all
  !
  performance-measurement
    delay-measurement
      profile test
    liveness-detection
      invalidation-action down
  !
!
```

### Configure PCE Node:

To configure PCE node:

```

pce
  address ipv4 9.9.9.9
  api
  !
  peer ipv4 1.1.1.1
  !
  peer ipv4 2.2.2.2
  !
  peer ipv4 4.4.4.4
  !
  segment-routing
    traffic-eng
      segment-list name srte11
        index 1 mpls adjacency 11.11.11.2
        index 2 mpls adjacency 13.13.13.2
        index 3 mpls adjacency 17.17.17.2
      !
      segment-list name srte12
        index 1 mpls adjacency 12.12.12.2
        index 2 mpls adjacency 15.15.15.2
        index 3 mpls adjacency 18.18.18.2
      !
      segment-list name srte13
        index 1 mpls adjacency 21.21.21.2
        index 2 mpls adjacency 22.22.22.2
        index 3 mpls adjacency 23.23.23.2
      !

  peer ipv4 1.1.1.1
  policy test
    color 10 end-point ipv4 2.2.2.2
    candidate-paths
      preference 100
        explicit segment-list srte11
      !
      !
      preference 200
        explicit segment-list srte13
      !
      !
      preference 300
        explicit segment-list srte12
      !
      !
      preference 400
        explicit segment-list srte11
```

```

!
!
!
profile-id 1
!

```

## Verification of SR-PCE: Enabling SR-PM Delay or Liveness for PCE-Initiated Policies

Use the **show segment-routing traffic-engineering policy all** command to verify the SR-PM delay or liveness for PCE-initiated policies configuration.

```

PE1(config)#do show segment-routing traffic-engineering policy all
Name: *2.2.2.2|10 (Color: 10 End-point: 2.2.2.2)
  Owners : PCEP
  Status:
    Admin: up, Operational: up for 13:50:38 (since 04-27 20:27:25.138)
  Candidate-paths:
    Preference 400 (PCEP):
      PM State: Up
      PCC profile: 1
      Dynamic (pce 9.9.9.9) (active)
        Metric Type: TE, Path Accumulated Metric: 0
          37 [Adjacency-SID, 11.11.11.1 - 11.11.11.2]
          28 [Adjacency-SID, 13.13.13.1 - 13.13.13.2]
          48 [Adjacency-SID, 17.17.17.1 - 17.17.17.2]
  Attributes:
    Binding SID: 153
    Allocation mode: dynamic
    State: Programmed
  Autoroute:
    Include all (Strict)

```

# Telemetry (Model-Based Telemetry and Event-Based Telemetry) Support for Performance Measurement

Table 7: Feature History

Feature Name	Release Information	Description
Telemetry (Model-Based Telemetry and Event-Based Telemetry) Support for Performance Measurement	Cisco IOS XE Amsterdam 17.3.1	<p>This feature enables Model-Based Telemetry (MDT) and Event-Based Telemetry (EDT) that allow the data to be directed to a configured receiver. This data can be used for analysis and troubleshooting purposes to maintain the health of the network.</p> <p>This feature is supported on Cisco ASR 900 RSP3 module.</p> <p>The <b>sr_5_label_push_enable</b> SDM template is mandatory for this feature to function.</p>

Telemetry is the process of measuring the state of the components in a system and transmitting it to a remote location for further processing and analysis.

The demand for data regarding network state, whether to detect hot spots in the network, or to aid decision making on workload placement requires data at a cadence that traditional methods cannot deliver. SNMP, CLI, and Syslog have limitations that restrict automation and scale.

Streaming telemetry lets users direct data to a configured receiver. This data can be used for analysis and troubleshooting purposes to maintain the health of the network. This is achieved by leveraging the capabilities of machine-to-machine communication.

Model-Driven Telemetry (MDT) is an approach for network monitoring in which data is streamed from the network devices continuously using a push model and provides near real-time access to operational statistics. Applications can subscribe to specific data items they need, by using standard-based YANG data models over NETCONF-YANG.



**Note** The **sr\_5\_label\_push\_enable** SDM template is mandatory for this feature to function.

## Probe, Aggregation, and Advertisement

Probe is a packet sent over a regular interval (probe interval) that carries the information about measurement (for example, delay, loss, and so on). The two types of probes are query and responder.

Aggregation is the process of aggregating the measurement values of the number of probes. The aggregation process is performed at a regular interval of time. The aggregation interval is usually a multiple of the probe interval; however, it can be as less as a probe interval.



Advertisement is a process of advertising the aggregated values when the measurement values cross the pre-determined threshold values. The advertisement check is performed after every aggregation interval. When the accelerated advertisement is configured, the check is performed in every probe interval.

### Configuration Methods of MDT

- **Cadence-Based Telemetry:** Cadence-based Telemetry (CDT) continuously streams data (operational statistics and state transitions) at a configured cadence. The streamed data helps users closely identify patterns in the networks (for example, streaming data about interface counters, and so on). Configuring the interval to any nonzero value sets the subscription for cadence-based telemetry.

It supports the Histogram Data. Histograms are more complex data type requiring the most processing on a device. Histograms store the frequency of occurrence over a time period and typically use ranges to group similar values. Histogram data provides the following information:

- Data of the history of the probe, aggregation, and advertisement
- Data for the last probe, last aggregation, and the last advertisement

- **Event-Based Telemetry:** Event-driven Telemetry (EDT) optimizes data collected at the receiver by streaming data only when a state transition occurs (for example, stream data only when an interface state transitions, IP route updates, and so on). Configuring the sample interval value to zero sets the subscription for event-based telemetry. EDT provides the following information:

- Delay metrics computed in the last probe-interval (Event: probe-completed)
- Delay metrics computed in the last aggregation-interval or the end of the periodic advertisement-interval (Event: advertisement-interval expired)
- Delay metrics last flooded in the network (Event: flooding-triggered)

The table below shows the data supported for link delay and end-to-end delay measurement in Oper-model. Oper-model is one of the categories in YANG model testing, where the operation data is pulled from the node.

Performance Measurement	Data Supported
Link Delay Measurement	Interface-last-probes
	interface-last-aggregations
	interface-last-advertisements
	interface-probe-histories
	interface-aggregated-histories
	interface-advertisement-histories

Performance Measurement	Data Supported
End-to-End Delay Measurement	sr-policy-last-probes
	sr-policy-last-aggregations
	sr-policy-last-advertisements
	sr-policy-probe-histories
	sr-policy-aggregated-histories
	sr-policy-advertisement-histories

For more information on the Telemetry feature, see the [Programmability Configuration Guide, Cisco IOS XE Amsterdam 17.1.x](#).

## Configuration Example: Telemetry for Performance Measurement

The following example shows the configuration example of telemetry for performance measurement (End-to-End Delay measurement) for the interface last advertisement option:

```
configure terminal
telemetry ietf subscription 100
encoding encode-kvgpb
filter xpath /performance-measurement/if-delay/last-advertisement
source-address <management-ip-address>
source-vrf management-interface
stream yang-push
update-policy periodic 100
receiver ip address <x.x.x.x> 57344 protocol grpc-tcp
```

The following example shows the sample output of telemetry configuration:

```
Node : <Router>
Subscription : 100
Path :
Cisco-IOS-XE-performance-measurement-oper:performance-measurement/if-delay/last-advertisement

Key : /if-name : GigabitEthernet0/0/13

/values/avg : 130
/values/min : 106
/values/max : 197
/values/var : 24
/timestamp : 2020-07-28T09:32:44+00:00
/advertised-reason : per-threshold-min
```

The options to configure telemetry performance measurement (Link Delay measurement) are:

- if-name /performance-measurement/if-delay/if-name
- probe is valid /performance-measurement/if-delay/probe-is-valid
- aggr is valid /performance-measurement/if-delay/aggr-is-valid
- adv is valid /performance-measurement/if-delay/adv-is-valid
- last probe /performance-measurement/if-delay/last-probe

- last aggr /performance-measurement/if-delay/last-aggr
- last adv /performance-measurement/if-delay/last-adv
- probe history /performance-measurement/if-delay/probe-history
- aggr-history /performance-measurement/if-delay/aggr-history
- adv history /performance-measurement/if-delay/adv-history

The options to configure telemetry performance measurement (End-to-End Delay measurement) are:

- sr-policy name /performance-measurement/sr-policy-delay/sr-policy-name
- sr-policy probe is valid /performance-measurement/sr-policy-delay/probe-is-valid
- aggr-is-vaidd /performance-measurement/sr-policy-delay/aggr-is-valid
- adv-is-valid /performance-measurement/sr-policy-delay/adv-is-valid
- lastprobe /performance-measurement/sr-policy-delay/last-probe
- probe history /performance-measurement/sr-policy-delay/probe-history
- last-aggr /performance-measurement/sr-policy-delay/last-aggr
- aggr-history /performance-measurement/sr-policy-delay/aggr-history
- last-adv /performance-measurement/sr-policy-delay/last-adv
- adv history /performance-measurement/sr-policy-delay/adv-history

## Verification of MDT and EDT Support for Performance Measurement

Use the following commands to verify the configuration of MDT and EDT for performance measurement.

```
Router#show performance-measurement history sr-policy probe
SR Policy name: foo
Candidate-Path:
  Preference           : 10
  Protocol-origin      : Configured
  Discriminator        : 0
  Active               : Yes
  Probe Start Timestamp Pkt(TX/RX) Average   Min      Max
    09:59:35  12 2020 3/3      303000  303000  303000
    09:59:30  12 2020 3/3      303000  303000  303000
    09:59:25  12 2020 3/3      302333  302000  303000
    09:59:20  12 2020 3/3      303000  303000  303000
    09:59:15  12 2020 3/3      303000  303000  303000
    09:59:10  12 2020 3/3      303000  303000  303000
    09:59:05  12 2020 3/3      302333  302000  303000
    09:59:00  12 2020 3/3      302333  302000  303000
    09:58:55  12 2020 3/3      303333  303000  304000
    09:58:50  12 2020 3/3      303000  303000  303000
    09:58:45  12 2020 3/3      302000  302000  302000
Segment-list:
  Name                  : SegmentList0
  Probe Start Timestamp Pkt(TX/RX) Average   Min      Max
    09:59:35  12 2020 3/3      303000  303000  303000
    09:59:30  12 2020 3/3      303000  303000  303000
    09:59:25  12 2020 3/3      302333  302000  303000
```

```

09:59:20 12 2020 3/3      303000      303000      303000
09:59:15 12 2020 3/3      303000      303000      303000
09:59:10 12 2020 3/3      303000      303000      303000
09:59:05 12 2020 3/3      302333      302000      303000
09:59:00 12 2020 3/3      302333      302000      303000
09:58:55 12 2020 3/3      303333      303000      304000
09:58:50 12 2020 3/3      303000      303000      303000
09:58:45 12 2020 3/3      302000      302000      302000
Atomic path:
Hops      : 192.168.0.2, 192.168.0.9
Labels    : 16151
Outgoing Interface : Ethernet0/1
Next Hop  : 11.11.11.2
Destination : 192.168.0.9
Session ID : 4
  Probe Start Timestamp Pkt(TX/RX) Average   Min      Max
    09:59:35 12 2020 1/1      303000      303000      303000
    09:59:30 12 2020 1/1      303000      303000      303000
    09:59:25 12 2020 1/1      303000      303000      303000
    09:59:20 12 2020 1/1      303000      303000      303000
    09:59:15 12 2020 1/1      303000      303000      303000
    09:59:10 12 2020 1/1      303000      303000      303000
    09:59:05 12 2020 1/1      303000      303000      303000
    09:59:00 12 2020 1/1      303000      303000      303000
    09:58:55 12 2020 1/1      304000      304000      304000
    09:58:50 12 2020 1/1      303000      303000      303000
    09:58:45 12 2020 1/1      302000      302000      302000
Atomic path:
Hops      : 192.168.0.2, 192.168.0.9
Labels    : 16151
Outgoing Interface : Ethernet0/2
Next Hop  : 12.12.12.2
Destination : 192.168.0.9
Session ID : 5
  Probe Start Timestamp Pkt(TX/RX) Average   Min      Max
    09:59:35 12 2020 1/1      303000      303000      303000
    09:59:30 12 2020 1/1      303000      303000      303000
    09:59:25 12 2020 1/1      302000      302000      302000
    09:59:20 12 2020 1/1      303000      303000      303000
    09:59:15 12 2020 1/1      303000      303000      303000
    09:59:10 12 2020 1/1      303000      303000      303000
    09:59:05 12 2020 1/1      302000      302000      302000
    09:59:00 12 2020 1/1      302000      302000      302000
    09:58:55 12 2020 1/1      303000      303000      303000
    09:58:50 12 2020 1/1      303000      303000      303000
    09:58:45 12 2020 1/1      302000      302000      302000
Atomic path:
Hops      : 192.168.0.2, 192.168.0.9
Labels    : 16151
Outgoing Interface : Ethernet0/3
Next Hop  : 13.13.13.2
Destination : 192.168.0.9
Session ID : 6
  Probe Start Timestamp Pkt(TX/RX) Average   Min      Max
    09:59:35 12 2020 1/1      303000      303000      303000
    09:59:30 12 2020 1/1      303000      303000      303000
    09:59:25 12 2020 1/1      302000      302000      302000
    09:59:20 12 2020 1/1      303000      303000      303000
    09:59:15 12 2020 1/1      303000      303000      303000
    09:59:10 12 2020 1/1      303000      303000      303000
    09:59:05 12 2020 1/1      302000      302000      302000
    09:59:00 12 2020 1/1      302000      302000      302000
    09:58:55 12 2020 1/1      303000      303000      303000

```

```

09:58:50 12 2020 1/1 303000 303000 303000
09:58:45 12 2020 1/1 302000 302000 302000

```

Router#show performance-measurement history sr-policy aggregation

SR Policy name: foo

Candidate-Path:

```

Preference           : 10
Protocol-origin      : Configured
Discriminator        : 0
Active               : Yes

```

Aggregation Timestamp	Average	Min	Max	Action
09:59:12 12 2020	302666	302000	304000	FIRST

Segment-list:

Name : SegmentList0

Aggregation Timestamp	Average	Min	Max	Action
09:59:12 12 2020	302666	302000	304000	FIRST

Atomic path:

```

Hops                  : 192.168.0.2, 192.168.0.9
Labels                : 16151
Outgoing Interface    : Ethernet0/1
Next Hop              : 11.11.11.2
Destination           : 192.168.0.9
Session ID            : 4

```

Aggregation Timestamp	Average	Min	Max	Action
09:59:12 12 2020	303000	302000	304000	FIRST

Atomic path:

```

Hops                  : 192.168.0.2, 192.168.0.9
Labels                : 16151
Outgoing Interface    : Ethernet0/2
Next Hop              : 12.12.12.2
Destination           : 192.168.0.9
Session ID            : 5

```

Aggregation Timestamp	Average	Min	Max	Action
09:59:12 12 2020	302499	302000	303000	FIRST

Atomic path:

```

Hops                  : 192.168.0.2, 192.168.0.9
Labels                : 16151
Outgoing Interface    : Ethernet0/3
Next Hop              : 13.13.13.2
Destination           : 192.168.0.9
Session ID            : 6

```

Aggregation Timestamp	Average	Min	Max	Action
09:59:12 12 2020	302499	302000	303000	FIRST

Router#show performance-measurement history sr-policy advertisement

SR Policy name: foo

Candidate-Path:

```

Preference           : 10
Protocol-origin      : Configured
Discriminator        : 0
Active               : Yes

```

Advertisement Timestamp	Average	Min	Max	Action
09:59:12 12 2020	302666	302000	304000	FIRST

Segment-list:

Name : SegmentList0

Advertisement Timestamp	Average	Min	Max	Action
09:59:12 12 2020	302666	302000	304000	FIRST

Atomic path:

```

Hops                  : 192.168.0.2, 192.168.0.9
Labels                : 16151
Outgoing Interface    : Ethernet0/1
Next Hop              : 11.11.11.2
Destination           : 192.168.0.9
Session ID            : 4

```

```

      Advertisement Timestamp Average   Min      Max      Action
      09:59:12  12 2020 303000      302000      304000      FIRST
Atomic path:
  Hops              : 192.168.0.2, 192.168.0.9
  Labels            : 16151
  Outgoing Interface : Ethernet0/2
  Next Hop          : 12.12.12.2
  Destination       : 192.168.0.9
  Session ID        : 5
      Advertisement Timestamp Average   Min      Max      Action
      09:59:12  12 2020 302499      302000      303000      FIRST
Atomic path:
  Hops              : 192.168.0.2, 192.168.0.9
  Labels            : 16151
  Outgoing Interface : Ethernet0/3
  Next Hop          : 13.13.13.2
  Destination       : 192.168.0.9
  Session ID        : 6
      Advertisement Timestamp Average   Min      Max      Action
      09:59:12  12 2020 302499      302000      303000      FIRST

Router#show performance-measurement history interfaces advertisement
Interface Name: Ethernet0/1 (ifh: 0x3)
Delay-Measurement history (uSec):
      Advertisement Timestamp Average   Min      Max      Action
      10:10:41  12 2020 204600      1          329999      FIRST

Router#show performance-measurement history interfaces aggregation
Interface Name: Ethernet0/1 (ifh: 0x3)
Delay-Measurement history (uSec):
      Aggregation Timestamp Average   Min      Max      Action
      10:10:41  12 2020 189405      1          329999      FIRST

Router#show performance-measurement history interfaces probe
Interface Name: Ethernet0/1 (ifh: 0x3)
Delay-Measurement history (uSec):
      Probe Start Timestamp Pkt(TX/RX) Average   Min      Max
      10:10:45  12 2020 3/3      202666      202499      202999
      10:10:35  12 2020 3/3      202999      202999      202999
      10:10:25  12 2020 3/3      202999      202999      202999
      10:10:15  12 2020 3/3      202333      201999      202500
      10:10:05  12 2020 3/3      203166      202999      203499
      10:09:55  12 2020 3/3      202999      202999      202999
      10:09:45  12 2020 3/3      249999      209999      329999
      10:09:35  12 2020 3/3      319999      309999      329999
      10:09:24  12 2020 3/3      326666      319999      329999
      10:09:14  12 2020 3/3      171499      499        299999
      10:09:04  12 2020 3/3      499        499        499
      10:08:54  12 2020 3/3      333        1          499

```

Use the command below to verify telemetry for End-to-End Delay Measurement:

```

Router#show performance-measurement history interfaces name gigabitEthernet <> advertisement
Interface Name: <> (ifh: 0x14)
Delay-Measurement history (uSec):
      Advertisement Timestamp Average   Min      Max      Action
      09:56:00  21 2020 161          100      462      PER-MIN

```

## Configuring UDP Destination Port

When you specify PM-UDP protocol, you need to configure the UDP destination port. The UDP port is configured for each PM measurement probe type (delay, loss, protocol, authentication mode, etc.) on querier

and responder nodes. The UDP port for each PM measurement probe type must match on querier and responder nodes.



---

**Note** The same UDP destination port is used for delay measurement for links and SR Policy.

---

This example shows how to configure the UDP destination port.

```
R1(config)#performance-measurement
R1(config-perf-meas)#protocol twamp-light
R1(config-pm-twamp)#measurement delay
R1(config-pm-twamp-delay)#unauthenticated
R1(config-pm-twamp-delay-unauth)#querier-dst-port 11222 querier-src-port 11333
R1(config-pm-twamp-delay-unauth)#exit
```







## CHAPTER 15

# Static Route Traffic Steering Using SR-TE Policy

Table 8: Feature History

Feature Name	Release Information	Description
Static Route Traffic Steering Using SR-TE Policy	Cisco IOS XE Amsterdam 17.3.1	<p>This feature allows the non colored (BGP Extended Community) prefix to steer traffic over static policy. Prior to this release, only colored (BGP Extended Community) prefix could automatically steer traffic based on the defined policy using a tunnel interface. Unlike non colored prefix, this was possible only for the colored prefix as it could match the SR policy.</p> <p>IPv4 static routes are now enhanced to leverage the SR policies to aid Segment Routing Traffic Engineering (SR-TE). This facilitates traffic steering for non colored prefix as you can now configure IP Static Route with SR static policy.</p> <p>The following new keyword for the <b>ip route</b> command is introduced:</p> <p><b>segment-routing policy</b> [<i>policy name</i>]</p>

Segment Routing (SR) allows a headend node to steer a packet flow along any path. This is accomplished by steering a flow into a SR policy. The header of the packet is then augmented with an ordered list of segments associated with the given SR policy. Thus, when colored (BGP Extended Community) prefix is received and there is a matching SR policy (ODN or Static), traffic for that prefix gets automatically steered based on the defined policy using a tunnel interface.

But, for non colored prefix, traffic is steered through the shortest IGP path. As a result, you cannot steer traffic for non colored prefix over static policy. Starting from Cisco IOS XE Release 17.3.1, IPv4 static routes are

enhanced to leverage the SR policies to aid Segment Routing Traffic Engineering (SR-TE). This facilitates traffic steering for non colored prefix as you can now configure IP Static Route with SR static policy.

**Note**

This feature is enabled by default; you do not require SDM template to enable this feature.

But, for Cisco ASR 900 RSP3 module, you must use **sr\_5\_label\_push\_enable** SDM template, if the number of SR labels is more.

- [Restrictions, on page 172](#)
- [Configuration Example: Static Route Traffic Steering Using SR-TE Policy, on page 172](#)
- [Verification of Static Route Traffic Steering Using SR-TE Configuration, on page 173](#)

## Restrictions

- VRF IP Static route is *not* supported.
- L2VPN traffic steering is *not* supported.
- IPv6 static route configuration is *not* supported.

## Configuration Example: Static Route Traffic Steering Using SR-TE Policy

The following example shows how to configure static route traffic steering using SR-TE policy:

```
enable
configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Router(config)#segment-routing mpls
Router(config-srmppls)# set-attributes
Router(config-srmppls-attr)# address-family ipv4
Router(config-srmppls-attr-af)# sr-label-preferred
Router(config-srmppls-attr-af)# explicit-null
Router(config-srmppls-attr-af)# exit-address-family
Router(config-srmppls-attr)# connected-prefix-sid-map
Router(config-srmppls-conn)# address-family ipv4
Router(config-srmppls-conn-af)# 1.1.1.1/32 index 1 range 1
Router(config-srmppls-conn-af)# exit-address-family
Router(config-srmppls-conn)#segment-routing traffic-eng
Router(config-srte)# segment-list name <segment-list name>
Router(config-srte-ep)# segment-list name <segment-list name>
Router(config-srte-ep)# index 1 mpls label 16005
Router(config-srte-ep)# index 2 mpls label 16010
Router(config-srte-ep)# index 3 mpls label 16009
Router(config-srte-ep)# policy <policy name>
Router(config-srte-policy)# color 50 end-point 21.21.21.21
Router(config-srte-policy)# candidate-paths
Router(config-srte-policy-path)# preference 100
Router(config-srte-policy-path-pref)# explicit segment-list <segment-list name>
Router(config-srte-policy-path-pref)# constraints
Router(config-srte-policy-path-pref-constr)# segments
Router(config-srte-policy-path-pref-constr-seg)# dataplane mpls
```

```
Router(config-srte-policy-path-pref-constr-seg)#end

Router#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Router(config)#ip route 21.21.21.21 255.255.255.255 segment-routing policy <policy name>
#
```

## Verification of Static Route Traffic Steering Using SR-TE Configuration

Use the following commands to verify static route traffic steering using SR-TE configuration.

```
Router#show traffic-eng policy all

Name: <policy name> (Color: 50 End-point: 21.21.21.21)
Owners : CLI
Status:
Admin: up, Operational: up for 24:27:47 (since 04-21 11:22:13.731)
Candidate-paths:
Preference 100 (CLI):
Explicit: segment-list Inter_IGP (active)
Weight: 1, Metric Type: TE
16005
16010
16009
Attributes:
Binding SID: 16
Allocation mode: dynamic
State: Programmed

Router#show traffic-eng policy name detail

Name: <policy name> (Color: 50 End-point: 21.21.21.21)
Owners : CLI
Status:
Admin: up, Operational: up for 24:28:01 (since 04-21 11:22:13.731)
Candidate-paths:
Preference 100 (CLI):
Explicit: segment-list Inter_IGP (active)
Weight: 1, Metric Type: TE
16005
16010
16009
Attributes:
Binding SID: 16
Allocation mode: dynamic
State: Programmed
Forwarding-ID: 65536 (0x30)
Per owner configs:
CLI
Binding SID: dynamic
Stats:
Packets: 0 Bytes: 0

Event history:
Timestamp Client Event type Context: Value
-----:-----
04-21 07:24:13.219 CLI Policy created Name: CLI
04-21 07:24:13.220 CLI Set colour Colour: 50
04-21 07:24:13.220 CLI Set end point End-point: 21.21.21.21
04-21 07:24:13.221 CLI Set explicit path Path option: Inter_IGP
```

```

04-21 07:24:16.357 CLI BSID allocated FWD: label 16
04-21 07:27:15.115 FH Resolution Policy state UP Status: PATH RESOLVED
04-21 07:27:47.305 FH Resolution REOPT triggered Status: REOPTIMIZED
04-21 07:39:11.343 FH Resolution REOPT triggered Status: REOPTIMIZED
04-21 07:39:23.939 FH Resolution REOPT triggered Status: REOPTIMIZED
04-21 11:09:12.319 FH Resolution REOPT triggered Status: REOPTIMIZED
04-21 11:22:13.580 FH Resolution REOPT triggered Status: REOPTIMIZED
04-21 11:22:13.731 FH Resolution Policy state DOWN Status: PATH NOT RESOLVED
04-21 11:37:05.691 FH Resolution Policy state UP Status: PATH RESOLVED
04-21 11:37:05.708 FH Resolution REOPT triggered Status: REOPTIMIZED
04-21 11:38:12.652 FH Resolution REOPT triggered Status: REOPTIMIZED
04-21 11:46:12.113 FH Resolution REOPT triggered Status: REOPTIMIZED
04-22 11:47:18.555 CLI Set explicit path Path option: Inter_IGP

Router#show ip route 21.21.21.21
Routing entry for 21.21.21.21/32
Known via "bgp 1", distance 200, metric 0, type internal
Last update from 6.6.6.6 1d00h ago
SR Incoming Label: 16009
Routing Descriptor Blocks:
* 6.6.6.6, from 6.6.6.6, 1d00h ago, merge-labels
opaque_ptr 0x3643B444
Route metric is 0, traffic share count is 1
AS Hops 0
MPLS label: 16009

Router#show ip cef 12.12.12.12

12.12.12.12/32
nexthop 192.168.1.2 GigabitEthernet0/0/0 label [16005|16005]-(local:16005) 16009-(local:16009)

##### Adding static route CLI herer #####
Router(config)#ip route 21.21.21.21 255.255.255.255 segment-routing policy <policy name>

Router#show ip route 21.21.21.21
Routing entry for 21.21.21.21/32
Known via "static", distance 1, metric 0
Routing Descriptor Blocks:
* Binding Label: 16
Route metric is 0, traffic share count is 1
MPLS label: implicit-null
Router#show ip cef 21.21.21.21
21.21.21.21/32
nexthop 16 Tunnel65536

```



## CHAPTER 16

# EVPN Single-Homing Over Segment Routing

Table 9: Feature History

Feature Name	Release Information	Feature Description
EVPN Single-Homing Over Segment Routing	Cisco IOS XE Amsterdam 17.3.1	<p>The EVPN Single-Homing feature utilizes the BGP MPLS-based Ethernet VPN functionality as defined in RFC 7432. That is, to achieve single-homing between a Provider Edge (PE) and a Customer Edge (CE) device.</p> <p>There are three fundamental building blocks for EVPN technology, EVPN Instance (EVI), Ethernet Segment (ES), EVPN BGP routes and extended communities.</p> <p>For EVPN Single-Homing feature, a CE device is attached to a single PE device and has an Ethernet Segment.</p>

The EVPN Single-Homing feature utilizes the functionality defined in RFC 7432 (BGP MPLS-based Ethernet VPN), to achieve single-homing between a Provider Edge (PE) and a Customer Edge (CE) device.

- [Information about EVPN Single-Homing, on page 176](#)
- [Prerequisites for EVPN Single-Homing, on page 180](#)
- [Restrictions for EVPN Single-Homing, on page 180](#)
- [How to Configure EVPN Single Homing, on page 181](#)
- [Verification Examples for EVPN Single-Homing, on page 185](#)
- [Additional References for EVPN Single-Homing, on page 191](#)

# Information about EVPN Single-Homing

## Ethernet Multipoint Connectivity

To achieve Ethernet multipoint connectivity, MPLS deployments traditionally rely on Virtual Private LAN Services (VPLS). A VPLS service is built with a full-mesh of pseudowires between PE devices that are part of a Layer 2 broadcast domain. A VPLS PE device performs data-plane MAC learning. For MAC learning, the VPLS PE device uses local interfaces for traffic coming from the access network and uses pseudowires for the traffic coming from the core network.

## EVPN Multipoint Solution

EVPN is the next generation of multipoint L2VPN solution that aligns operation principles of L3VPN with Ethernet services. Instead of relying solely on data plane for MAC Address learning, EVPN PE devices signal and learn MAC addresses over the core network using BGP, while still using data plane MAC-learning on the access side. Providers can configure BGP as a common VPN control plane for their ethernet offerings and leverage the advantages of Layer 3 VPN over VPLS.

## EVPN Building Blocks

There are three fundamental building blocks for EVPN technology, EVPN Instance (EVI), Ethernet Segment (ES), EVPN BGP routes and extended communities:

- EVI is a VPN connection on a PE router. It is the equivalent of IP VPN Routing and Forwarding (VRF) in Layer 3 VPN. It is also known as MAC-VRF.
- ES is a connection with a customer site (device or network) and is associated with access-facing interfaces. Access-facing interfaces are assigned unique IDs that are referred to as Ethernet Segment Identifiers (ESI). A site can be connected to one or more PEs. The ES connection has the same ESI in each PE connected to the site.
- RFC 7432 defines routes and extended communities to enable EVPN support. In Cisco IOS XE Fuji 16.8.x Software Release, Route Type 2 and Route Type 3 are supported.

In BGP MPLS-based EVPN, an EVI is configured for every PE device for each customer associated with the PE device. In this case, a customer is any customer edge device that is attached to the PE device. The CE device can be a host, a switch or a router. Each EVI has a unique Route Distinguisher (RD) and one or more Route Targets (RT).

For EVPN Single-Homing feature, a CE device is attached to a single PE device and has an Ethernet Segment with ESI=0.

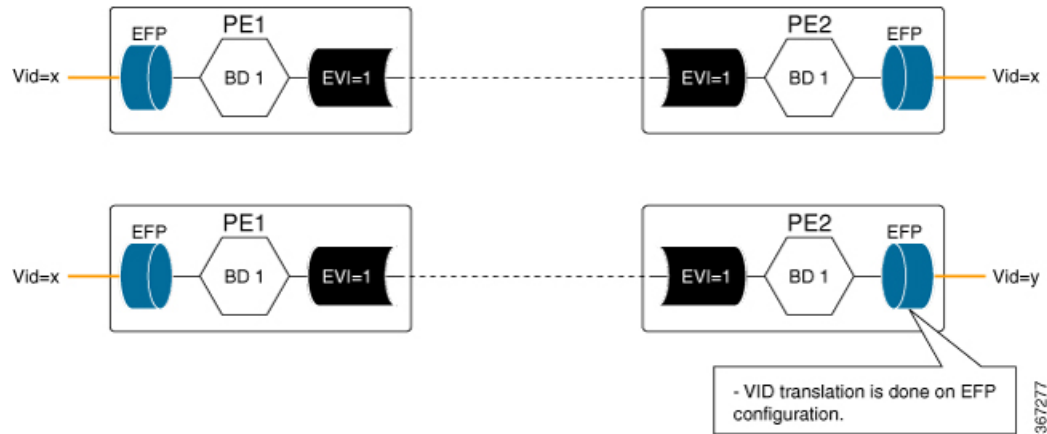
## Service Interfaces

The following are types of EVPN VLAN service interfaces:

### VLAN-based Service Interface

In VLAN-based service interface, each VLAN is associated to one bridge domain and one EVI.

Figure 19: VLAN-Based Service Interface



For VLAN-based Service Interface, Type 1 Route Distinguisher, a unique number used to distinguish identical routes in different VRFs, is used for EVIs as recommended by the RFC 7432. The Route Distinguishers and Router Targets, which are used to share routes between different VRFs, are autogenerated to ensure unique Route Distinguisher numbers across EVIs.

### VLAN Bundle Service Interface

In VLAN Bundle Service Interface, multiple VLANs share the same bridge table.

Figure 20: VLAN Bundle Service Interface

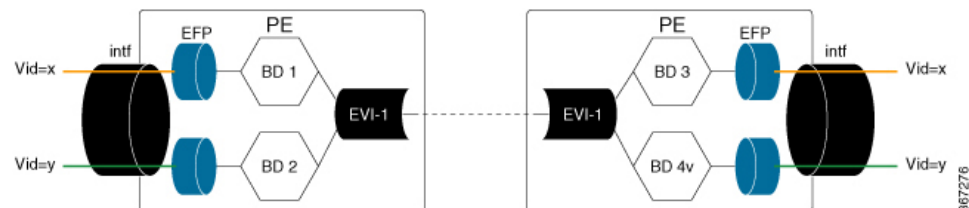


Each EVPN instance corresponds to multiple broadcast domains maintained in a single bridge table per MAC-VRF. For VLAN Bundle Service Interface service to work, MAC addresses must be unique across all VLANs for an EVI.

### VLAN-Aware Bundle Service Interface

For VLAN-aware Bundle Service Interface, each VLAN is associated with one bridge domain, but there can be multiple bridge domains associated with one EVI.

Figure 21: VLAN-Aware Bundle Service Interface



An EVPN instance consists of multiple broadcast domains where each VLAN has one bridge table. Multiple bridge tables (one per VLAN) are maintained by a single MAC-VRF that corresponds to the EVPN instance.

## Route Types

For EVPN Single-Homing feature, Route Type 2 and Route Type 3 are supported, as defined by RFC 7432.

### Route Type 2 — MAC and IP Advertisement Route

Type 2 Routes are used to advertise MAC addresses and their associated IP addresses. When a PE router learns the MAC address of a CE device that is connected to it locally, or a MAC address of a device behind the CE device, a MAC and an IP advertisement route is created.

The following table describes the header format for the MAC and IP Advertisement Route packet:

**Table 10: Header format for the MAC and IP Advertisement Route packet**

Field	Value	Length (Octets)
Route Type	0x02	1
Length	Variable	1
EVI RD	Type 1 (IPv4 address) RD unique across all EVIs on the PE	8
ESI	Ethernet Segment Identifier	10
Ethernet Tag	0 or valid Ethernet Tag	4
MAC Addr Len	48	1
MAC Address	Valid MAC address	6
IP Addr Length	IP address length in bits: 0, 32 or 128	1
IP Address	Optional IP address	0 or 4 or 16
Label1	Valid downstream assigned label to perform forwarding to a CE device based on the destination MAC address	3
Label2	Specifies a second label	0-3
EVI RT	Type 0 (2byteAS) route target	8
MAC Mobility	0x0600:{ 1 byte Sticky bit}:0x00:{ 4 byte sequence number}	8



**Note**

- MAC Address field is populated with the CE address.
- IP address field is optional with IP Address length set to 0 bits.
- For EVPN Single-Homing feature, ESI value is always set to 0.
- In the Label field (Label1, Label2), Per-BD or Per-CE labels can be assigned.
  - Per-BD is used when PE advertises a single label for all MAC addresses learned in a given bridge domain.
  - Per-CE label assigns a separate label to each access port in the bridge domain.

**Route Type 3 — Inclusive Multicast Ethernet Tag Route**

Type 3 routes are used for transporting Broadcast, Unknown Unicast, and Multicast (BUM) traffic to other PE devices across a given EVPN network instance.

The following tables describes the header format for Type 3 routes:

**Table 11: Header Format for Type 3 Route Packets**

Field	Value	Length (Octets)
Route Type	0x03	1
Length	26 or 38	1
EVI RD	Type 1 (IPv4Addr) RD unique across all EVIs on the PE	8
Ethernet Tag	0 or valid Ethernet Tag	4
IP Addr Length	IP Address Length - 32 bits or 128 bits	1
IP Address	IP Address common for all EVIs (for example, loopback address)	4 or 16
PMSI Tunnel Attr	{ 1 byte flags = 0 }; { 1 byte Tunnel Type }; { 3 byte label }; { variable length Tunnel Identifier }	Variable
EVI RT	Type 0 (2byteAS) route target	8

The PE devices advertise an Inclusive Multicast Ethernet Tag (IMET) Route for every EVI-Ethernet Tag sequence. The Ethernet Tag is set to 0 for VLAN-based and VLAN-bundling service interfaces. The Ethernet Tag is set to a valid VLAN ID for VLAN-aware bundling service interface.

Type 3 route also carries a Provider Multicast Service Interface (PMSI) Tunnel attribute as specified in RFC 6514 (BGP Encodings and Procedures for MVPNs).

For Ingress Replication, the IMET route is used to advertise the label (in the PMSI Tunnel Attribute) that the other PEs can use to send BUM traffic to the originating PE device.

## Prerequisites for EVPN Single-Homing

- EVI and Bridge domains must be in established state with associated MPLS labels.

## Restrictions for EVPN Single-Homing

- Route Type 1 and Route Type 4 are not supported.
- Per-EVI-based labelling is not supported.
- Maximum number of supported bridge domains is 1600.
- Maximum number of supported EEPs or service instances is 8000.
- Single-Homing feature is not supported with port channel interface between Provider Edge and Customer Edge devices.
- ESI must be all 0s.
- BDI for EVPN bridge domain is not supported.
- EVPN feature does not work with VPLS template enabled.
- EVPN VC statistics not supported with default template.
- SR-TE is not supported with EVPN.
- MACSec with EVPN is not supported.
- EVPN MAC scale is limited to 20000 MACs at up to a maximum rate of 500 PPS.
- EVPN is not supported on TEFP.
- EVPN is not supported with "enable\_8k\_efp" template.
- Logical port IDs are allotted for Broadcast, Unicast, and Multicast (BUM) traffic and Unicast traffic. So, each EVPN session with a single neighbour utilizes two lportid from the pseudowire bucket and thus, is considered as two pseudowires.
- If the EVPN session neighbour increase, there is one more resource consumed for that EFI interface.
- For VPLS, single lportid per bridge-domain is allocated for each neighbour.





**Note** In the above example, the **l2vpn evpn instance** command and the associated sub-mode is only required if one or more of the following apply:

- There is per-EVI configuration to be applied (for example, route targets or route distinguisher)
- The EVI is VLAN-bundle or VLAN-aware.

If the EVPN instance is not explicitly configured, it is created automatically as a VLAN-based EVI with autogenerated route targets and route distinguisher.

### Configuring L2VPN EVPN Globally and EVI on IOS-XE Router

```
l2vpn evpn
  replication-type ingress      ----> Enables ingress replication label
!
l2vpn evpn instance 10 vlan-based ----> Configures Vlan-based EVI 10
!
l2vpn evpn instance 20 vlan-bundle ----> Configures Vlan-bundled EVI 20
!
l2vpn evpn instance 30 vlan-aware  ----> Configures Vlan-aware EVI 30
```

### Configuring Bridge Domains on IOS-XE Router

```
bridge-domain 10
  mac aging-time 30
  member GigabitEthernet6 service-instance 10 --> Links SI 10 on interface with Bridge-domain 10
  member evpn-instance 10 --> Links EVI 10 with Bridge-domain 10
!
bridge-domain 20
  mac aging-time 30
  member GigabitEthernet6 service-instance 20 --> Links SI 20 on interface with Bridge-domain 20
  member evpn-instance 20 --> Links EVI 20 with Bridge-domain 20
!
bridge-domain 30
  mac aging-time 30
  member GigabitEthernet6 service-instance 30 --> Links SI 30 on interface with Bridge-domain 30
  member evpn-instance 30 ethernet-tag 30 ----> Links EVI 30 with Bridge-domain 30
```

### Configuring Access Interface on a Provider Edge

```
interface GigabitEthernet6
  no ip address
  negotiation auto
  service instance 10 ethernet ----> Enables service instance 10 under the physical interface

  encapsulation dot1q 10
  !
  service instance 20 ethernet ----> Enables service instance 20 under the physical interface

  encapsulation dot1q 20-21
  !
  service instance 30 ethernet ----> Enables service instance 30 under the physical interface

  encapsulation dot1q 30
```

### Configuring Native SR for EVPN

```
segment-routing mpls

!

set-attributes

address-family ipv4

sr-label-preferred

exit-address-family

!

global-block 17000 23999

!

connected-prefix-sid-map

address-family ipv4

4.4.4.4/32 index 19 range 1

exit-address-family

!

router ospf 10

router-id 4.4.4.4

nsr

nsf cisco

segment-routing mpls

segment-routing prefix-sid-map advertise-local

fast-reroute per-prefix enable area 0 prefix-priority high

fast-reroute per-prefix enable prefix-priority low

fast-reroute per-prefix ti-lfa

microloop avoidance segment-routing

redistribute connected

network 0.0.0.0 255.255.255.255 area 0

bfd all-interfaces
```

## Configuring EVPN Single-Homing

Use the following steps to configure EVPN Single-Homing:

## Configuring BGP on Provider Edge Device, PE1

```
enable
configure terminal
router bgp 100
  bgp router-id 10.1.1.1
  bgp log-neighbor-changes
  bgp graceful-restart
  neighbor 10.2.2.2 remote-as 100
  neighbor 10.2.2.2 update-source Loopback0
!
address-family ipv4
  neighbor 10.2.2.2 activate
exit-address-family
!
address-family l2vpn evpn      ----> Enables L2VPN EVPN address family
  neighbor 10.2.2.2 activate
  neighbor 10.2.2.2 send-community both
  neighbor 10.2.2.2 soft-reconfiguration inbound
exit-address-family
```

## Configuring BGP on Route Reflector

```
router bgp 100
  bgp router-id 10.2.2.2
  bgp log-neighbor-changes
  bgp graceful-restart
  neighbor 10.1.1.1 remote-as 100
  neighbor 10.1.1.1 update-source Loopback0
  neighbor 10.3.3.3 remote-as 100
  neighbor 10.3.3.3 update-source Loopback0
  neighbor 10.7.7.7 remote-as 100
  neighbor 10.7.7.7 update-source Loopback0
!
address-family ipv4
  neighbor 10.1.1.1 activate
  neighbor 10.1.1.1 route-reflector-client
  neighbor 10.3.3.3 activate
  neighbor 10.3.3.3 route-reflector-client
  neighbor 10.7.7.7 activate
  neighbor 10.7.7.7 route-reflector-client
exit-address-family
!
address-family l2vpn evpn      ----> Enables L2vpn evpn address family
  neighbor 10.1.1.1 activate
  neighbor 10.1.1.1 send-community both
  neighbor 10.1.1.1 route-reflector-client
  neighbor 10.1.1.1 soft-reconfiguration inbound
  neighbor 10.3.3.3 activate
  neighbor 10.3.3.3 send-community both
  neighbor 10.3.3.3 route-reflector-client
  neighbor 10.3.3.3 soft-reconfiguration inbound
  neighbor 10.7.7.7 activate
  neighbor 10.7.7.7 send-community both
  neighbor 10.7.7.7 route-reflector-client
  neighbor 10.7.7.7 soft-reconfiguration inbound
exit-address-family
```

## Configuring Customer Edge and Provider Edge Interfaces

CE1 configuration

```

interface GigabitEthernet6.10
 encapsulation dot1Q 10
 ip address 203.0.113.1 255.255.255.240
interface GigabitEthernet6.20
 encapsulation dot1Q 20
 ip address 203.0.113.17 255.255.255.240
interface GigabitEthernet6.30
 encapsulation dot1Q 30
 ip address 203.0.113.33 255.255.255.240

```

### PE1 Configuration

```

interface GigabitEthernet6
 no ip address
 negotiation auto
 service instance 10 ethernet
 encapsulation dot1q 10
 !
 service instance 20 ethernet
 encapsulation dot1q 20-21
 !
 service instance 30 ethernet
 encapsulation dot1q 30

```

## Verification Examples for EVPN Single-Homing

Use the following command to verify that EVI and Bridge domains are in established state and to display associated MPLS labels:

### show l2vpn evpn evi detail

```

EVPN instance:      10 (VLAN Based) ----> VLAN Based EVI
RD:                 10.1.1.1:10 (auto) ----> RD derived from Loopback0 of PE1
Import-RTs:         100:10
Export-RTs:         100:10
Per-EVI Label:      none
State:              Established ----> EVI state
Encapsulation:      mpls
Bridge Domain:      10
Ethernet-Tag:       0
BUM Label:          23 ----> Broadcast/Unknown unicast/Multicast traffic label
Per-BD Label:       22
State:              Established ----> Bridge-domain state
Pseudoports:
  GigabitEthernet6 service instance 10 ----> Local interface part of bridge-domain
  GigabitEthernet7 service instance 10 ----> Local interface part of bridge-domain

EVPN instance:      20 (VLAN Bundle) ----> VLAN Bundled EVI
RD:                 10.1.1.1:20 (auto)
Import-RTs:         100:20
Export-RTs:         100:20
Per-EVI Label:      none
State:              Established
Encapsulation:      mpls
Bridge Domain:      20
Ethernet-Tag:       0
BUM Label:          20
Per-BD Label:       21
State:              Established
Pseudoports:
  GigabitEthernet6 service instance 20

```

```

GigabitEthernet7 service instance 20

EVPN instance:      30 (VLAN Aware) ----> VLAN-Aware EVI
RD:                 10.1.1.1:30 (auto)
Import-RTs:         100:30
Export-RTs:         100:30
Per-EVI Label:      none
State:              Established
Encapsulation:      mpls
Bridge Domain:      30
  Ethernet-Tag:      30
  BUM Label:         18
  Per-BD Label:      19
  State:             Established
Pseudoports:
  GigabitEthernet6 service instance 30
  GigabitEthernet7 service instance 30

```

Use the following command to see the L2VPN EVPN summary:

```

PE1#show l2vpn evpn summary

L2VPN EVPN

EVPN Instances (excluding point-to-point): 3

VLAN Aware: 1

VLAN Based: 1

VLAN Bundle: 1

Bridge Domains: 4

BGP: ASN 100, address-family l2vpn evpn configured

Router ID: 2.2.2.2

Label Allocation Mode: Per-BD

Global Replication Type: Ingress

MAC Duplication: seconds 180 limit 5

MAC Addresses: 6

Local: 3

Remote: 3

Duplicate: 0

CS1#

```

Use the following command to verify that the bridge domain has learnt the local and remote MAC addresses:

```

PE1#show bridge-domain 10
Bridge-domain 10 (3 ports in all)
State: UP                               Mac learning: Enabled
Aging-Timer: 30 second(s) ----> MAC aging timer for bridge-domain
  GigabitEthernet6 service instance 10

```



```
GigabitEthernet7 service instance 10
EVPN Instance 10
AED MAC address      Policy  Tag      Age  Pseudoport
- 000C.29B0.3E16 forward static_r  0    OCE_PTR:0xe8eb04a0 ---> Remotely learnt MAC
- 000C.29AF.F904 forward dynamic_c 29  GigabitEthernet6.EFP10 --> MAC locally learnt

- 000C.2993.130E forward dynamic_c 26  GigabitEthernet7.EFP10
- 000C.29EE.EC0D forward static_r  0    OCE_PTR:0xe8eb0500
```



**Note** In the above output, MAC addresses with forward dynamic\_c tags are locally learned addresses and MAC addresses with forward static\_r tags are remote addresses learned through EVPN.

Use the following command to verify that EVPN manager has received the local MACs learned by the bridge domain:

```
PE1# show l2vpn evpn mac
MAC Address      EVI    BD      ESI                      Ether Tag  Next Hop
-----
000c.2993.130e 10      10      0000.0000.0000.0000.0000  0          Gi7:10
000c.29af.f904 10      10      0000.0000.0000.0000.0000  0          Gi6:10
000c.29b0.3e16 10      10      0000.0000.0000.0000.0000  0          10.7.7.7
000c.29ee.ec0d 10      10      0000.0000.0000.0000.0000  0          10.3.3.3
```

```
PE1# show l2vpn evpn mac detail
MAC Address:          000c.2993.130e
EVPN Instance:        10
Bridge Domain:        10
Ethernet Segment:     0000.0000.0000.0000.0000
Ethernet Tag ID:      0
Next Hop(s):          GigabitEthernet7 service instance 10
Label:                22
Sequence Number:      0
MAC only present:     Yes
MAC Duplication Detection: Timer not running

MAC Address:          000c.29ee.ec0d
EVPN Instance:        10
Bridge Domain:        10
Ethernet Segment:     0000.0000.0000.0000.0000
Ethernet Tag ID:      0
Next Hop(s):          10.3.3.3
Local Address:        10.1.1.1
Label:                19
Sequence Number:      0
MAC only present:     Yes
MAC Duplication Detection: Timer not running
```



**Note** In the above output, the next hop address of the remote MAC is the address of the provider edge device, if it is learnt remotely or the local interface if MAC address is learnt locally.

Use the following command to verify that Layer 2 Routing Information Base (RIB) has the required the MAC info:

```

PE1# show l2route evpn mac
-----
EVI      ETag  Prod  Mac Address      Next Hop(s) Seq Number
-----
10        0 L2VPN 000C.2993.130E    Gi7:10          0
10        0 L2VPN 000C.29AF.F904    Gi6:10          0
10        0 BGP   000C.29B0.3E16    L:19 IP:10.7.7.7 0
10        0 BGP   000C.29EE.EC0D    L:19 IP:10.3.3.3 0

```

**Note**

Remote MACs are learnt through BGP. In the above command output, the producer is BGP and local MACs are learned through Layer 2 VPN.

Use the following command to verify that Layer 2 FIB has received the MAC information from Layer 2 RIB, and bridge-domain and MFI are configured.

```

PE1# show l2fib bridge-domain 10 detail
Bridge Domain : 10
Reference Count : 18
Replication ports count : 4
Unicast Address table size : 4
IP Multicast Prefix table size : 4

Flood List Information :
Olist: Id 9225, Port Count 4

Port Information :
Serv Inst: Gi6:10
Serv Inst: Gi7:10
EVPN MPLS Encap: pathlist 107
EVPN MPLS Encap: pathlist 101

Unicast Address table information :
Mac: 000c.2993.130e, Adjacency: Serv Inst: Gi7:10
Mac: 000c.29af.f904, Adjacency: Serv Inst: Gi6:10
Mac: 000c.29b0.3e16, Adjacency: EVPN MPLS Encap: pathlist 98
Mac: 000c.29ee.ec0d, Adjacency: EVPN MPLS Encap: pathlist 104

IP Multicast Prefix table information :
Source: *, Group: 224.0.0.0/4, IIF: , Adjacency: Olist: 9226, Ports: 0
Source: *, Group: 224.0.0.0/24, IIF: , Adjacency: Olist: 9225, Ports: 4
Source: *, Group: 224.0.1.39, IIF: , Adjacency: Olist: 9225, Ports: 4
Source: *, Group: 224.0.1.40, IIF: , Adjacency: Olist: 9225, Ports:

```

Use the following command to verify that the information on BGP route type 3 is sent to L2RIB:

```

PE1# show l2route evpn imet
-----
EVI      ETAG  Prod  Router IP Addr  Type  Label      Tunnel ID
-----
10        0 BGP   10.3.3.3      6     18         10.3.3.3
10        0 BGP   10.7.7.7      6     18         10.7.7.7
10        0 L2VPN 10.1.1.1      6     23         10.1.1.1

```

Use the following command to verify MPLS forwarding:

```

PE1# show mpls forwarding-table
Local      Outgoing  Prefix      Bytes Label  Outgoing  Next Hop
Label      Label     or Tunnel Id  Switched     interface

```

18	No Label	evpn(mc:bd 30)	305042	none	point2point
19	No Label	evpn(uc:bd 30)	7684	none	point2point
20	No Label	evpn(mc:bd 20)	542588	none	point2point
21	No Label	evpn(uc:bd 20)	13786	none	point2point
22	No Label	evpn(uc:bd 10)	6638	none	point2point
23	No Label	evpn(mc:bd 10)	277740	none	point2point
24	Pop Label	192.0.2.2-A	0	Gi1	192.0.2.2
25	Pop Label	192.0.2.2-A	0	Gi1	192.0.2.2
16001	16001	10.3.3.3/32	0	Gi1	192.0.2.2
16002	Pop Label	10.2.2.2/32	0	Gi1	192.0.2.2
16004	16004	10.7.7.7/32	0	Gi1	192.0.2.2

PE1# show ip bgp l2vpn evpn route-type 2

BGP routing table entry for [2][10.1.1.1:10][0][48][000C2993130E][0][\*]/20, version 43  
 Paths: (1 available, best #1, table evi\_10)

Advertised to update-groups:

2

Refresh Epoch 1

Local

:: (via default) from 0.0.0.0 (10.1.1.1)

Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best

EVPN ESI: 00000000000000000000, Label1 22

Extended Community: RT:100:10

rx pathid: 0, tx pathid: 0x0

BGP routing table entry for [2][10.1.1.1:10][0][48][000C29B03E16][0][\*]/20, version 116

Paths: (1 available, best #1, table evi\_10)

Not advertised to any peer

Refresh Epoch 3

Local, (received & used), imported path from [2][10.7.7.7:10][0][48][000C29B03E16][0][\*]/20  
 (global)

10.7.7.7 (metric 30) (via default) from 10.2.2.2 (10.2.2.2)

Origin incomplete, metric 0, localpref 100, valid, internal, best

EVPN ESI: 00000000000000000000, Label1 19

Extended Community: RT:100:10

Originator: 10.7.7.7, Cluster list: 10.2.2.2

rx pathid: 0, tx pathid: 0x0

BGP routing table entry for [2][10.1.1.1:10][0][48][000C29B03E16][0][\*]/20, version 116

Paths: (1 available, best #1, table evi\_10)

Not advertised to any peer

Refresh Epoch 3

Local, (received & used), imported path from [2][10.7.7.7:10][0][48][000C29B03E16][0][\*]/20  
 (global)

10.7.7.7 (metric 30) (via default) from 10.2.2.2 (10.2.2.2)

Origin incomplete, metric 0, localpref 100, valid, internal, best

EVPN ESI: 00000000000000000000, Label1 19

Extended Community: RT:100:10

Originator: 10.7.7.7, Cluster list: 10.2.2.2

rx pathid: 0, tx pathid: 0x0

BGP routing table entry for [2][10.1.1.1:10][0][48][000C29EEEC0D][0][\*]/20, version 134

Paths: (1 available, best #1, table evi\_10)

Not advertised to any peer

Refresh Epoch 3

Local, (received & used), imported path from [2][10.3.3.3:10][0][48][000C29EEEC0D][0][\*]/20  
 (global)

10.3.3.3 (metric 30) (via default) from 10.2.2.2 (10.2.2.2)

Origin incomplete, metric 0, localpref 100, valid, internal, best

EVPN ESI: 00000000000000000000, Label1 19

Extended Community: RT:100:10

Originator: 10.3.3.3, Cluster list: 10.2.2.2

rx pathid: 0, tx pathid: 0x0

PE1# show ip bgp l2vpn evpn route-type 3

BGP routing table entry for [3][10.1.1.1:10][0][32][10.1.1.1]/17, version 41

```

Paths: (1 available, best #1, table evi_10)
  Advertised to update-groups:
    2
  Refresh Epoch 1
  Local
    :: (via default) from 0.0.0.0 (10.1.1.1)
      Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
      Extended Community: RT:100:10
      PMSI Attribute: for EVPN, Flags: 0x0, Tunnel type: 6, length 4, label: 23 (vni 368)
  tunnel parameters: 0101 0101
    rx pathid: 0, tx pathid: 0x0
  BGP routing table entry for [3][10.1.1.1:10][0][32][10.3.3.3]/17, version 137
  Paths: (1 available, best #1, table evi_10)
    Not advertised to any peer
    Refresh Epoch 3
    Local, (received & used), imported path from [3][10.3.3.3:10][0][32][10.3.3.3]/17 (global)

      10.3.3.3 (metric 30) (via default) from 10.2.2.2 (10.2.2.2)
        Origin incomplete, metric 0, localpref 100, valid, internal, best
        Extended Community: RT:100:10
        Originator: 10.3.3.3, Cluster list: 10.2.2.2
        PMSI Attribute: for EVPN, Flags: 0x0, Tunnel type: 6, length 4, label: 18 (vni 288)
  tunnel parameters: 0303 0303
    rx pathid: 0, tx pathid: 0x0
  BGP routing table entry for [3][10.1.1.1:10][0][32][10.7.7.7]/17, version 122
  Paths: (1 available, best #1, table evi_10)
    Not advertised to any peer
    Refresh Epoch 3
    Local, (received & used), imported path from [3][10.7.7.7:10][0][32][10.7.7.7]/17 (global)

      10.7.7.7 (metric 30) (via default) from 10.2.2.2 (10.2.2.2)
        Origin incomplete, metric 0, localpref 100, valid, internal, best
        Extended Community: RT:100:10
        Originator: 10.7.7.7, Cluster list: 10.2.2.2
        PMSI Attribute: for EVPN, Flags: 0x0, Tunnel type: 6, length 4, label: 18 (vni 288)
  tunnel parameters: 0707 0707
    rx pathid: 0, tx pathid: 0x0

```

Use the following command to verify that segment routing table details:

```
PE1#show segment-routing mpls connected-prefix-sid-map ipv4
```

```
PREFIX_SID_CONN_MAP ALGO_0
```

Prefix/masklen	SID Type	Range	Flags	SRGB
1.1.1.1/32	18 Indx	1		Y

```
PREFIX_SID_PROTOCOL_ADV_MAP ALGO_0
```

Prefix/masklen	SID Type	Range	Flags	SRGB	Source
1.1.1.1/32	18 Indx	1		Y	IS-IS Level 1 0002.0000.0001
2.2.2.2/32	19 Indx	1		Y	IS-IS Level 1 0002.0000.0006
3.3.3.3/32	20 Indx	1		Y	IS-IS Level 1 0002.0000.0002
4.4.4.4/32	21 Indx	1		Y	IS-IS Level 1 0002.0000.0003

```
PE1#show segment-routing mpls state
```

```
Segment Routing MPLS State : ENABLED
```

## Additional References for EVPN Single-Homing

### Standards and RFCs

Standard	Title
RFC 7432	BGP MPLS-Based Ethernet VPN





## CHAPTER 17

# Enabling Segment Routing Flexible Algorithm

Table 12: Feature History

Feature Name	Release Information	Feature Description
IS-IS Flexible Algorithm Include Affinity Support	Cisco IOS XE Bengaluru 17.6.1	This feature supports "include-any" and "include-all" affinities in IS-IS. Prior to Cisco IOS XE Bengaluru 17.6.1 release, only Flexible Algorithm affinity "exclude-any" was supported.
Segment Routing Flexible Algorithm with OSPF	Cisco IOS XE Bengaluru 17.5.1	This feature allows you to configure Segment Routing Flexible Algorithm with OSPF. Flexible Algorithm with OSPF supports metric minimization and avoidance, multi-plane, delay metric with rounding, and ODN with auto-steering.

Feature Name	Release Information	Feature Description
Segment Routing Flexible Algorithm support for TI-LFA uLoop Avoidance, SID Leaking, and ODN with Auto-Steering	Cisco IOS XE Bengaluru 17.4.1	<p>This feature allows you to compute Loop Free Alternate (LFA) paths, TI-LFA backup paths, and Microloop Avoidance paths for a particular Flexible Algorithm using the same constraints as the calculation of the primary paths for such Flexible Algorithms, for IS-IS. See <a href="#">Calculation of Flexible Algorithm Path, on page 196</a></p> <p>Inter-area leaking of Flexible Algorithm SIDs and prefixes and selectively filtering the paths that are installed to the MFI are also supported. See <a href="#">Flexible Algorithm Prefix-SID Advertisement, on page 196</a> and <a href="#">Installation of Forwarding Entries for Flexible Algorithm Paths, on page 201</a></p>
Segment Routing Flexible Algorithm with IS-IS	Cisco IOS XE Amsterdam 17.3.1	<p>This feature allows you to customize IGP shortest path computation according to your needs. You can assign custom SR prefix-SIDs to forward the packets beyond link-cost-based SPF. As a result, a traffic engineered path is automatically computed by the IGP to any destination reachable by the IGP.</p>

Segment Routing Flexible Algorithm allows operators to customize IGP shortest path computation according to their own needs. An operator can assign custom SR prefix-SIDs to realize forwarding beyond link-cost-based SPF. As a result, Flexible Algorithm provides a traffic engineered path automatically computed by the IGP to any destination reachable by the IGP.

The SR architecture associates prefix-SIDs to an algorithm which defines how the path is computed. Flexible Algorithm allows for user-defined algorithms where the IGP computes paths based on a user-defined combination of metric type and constraint.

This document describes the IS-IS and OSPF extensions to support Segment Routing Flexible Algorithm on an MPLS data-plane.

- [Prerequisites for Flexible Algorithm, on page 195](#)
- [Restrictions for Flexible Algorithm, on page 195](#)
- [Building Blocks of Segment Routing Flexible Algorithm, on page 195](#)
- [Flexible Algorithm Configurations, on page 201](#)
- [Verifying the Flexible Algorithm Configuration, on page 207](#)



## Prerequisites for Flexible Algorithm

Segment routing must be enabled on the router before the Flexible Algorithm functionality is activated.

## Restrictions for Flexible Algorithm

- A maximum of four IS-IS flexible algorithm sessions are supported.
- In IS-IS, only the flexible algorithm affinity "exclude-any" is supported. "include-any" and "include-all" are not supported.

## Building Blocks of Segment Routing Flexible Algorithm

This section describes the building blocks that are required to support the SR Flexible Algorithm functionality in IS-IS and OSPF.

## Flexible Algorithm Definition

Many possible constraints may be used to compute a path over a network. Some networks are deployed with multiple planes. A simple form of constraint may be to use a particular plane. A more sophisticated form of constraint can include some extended metric, like delay, as described in [RFC7810]. Even more advanced case could be to restrict the path and avoid links with certain affinities. Combinations of these are also possible. To provide a maximum flexibility, the mapping between the algorithm value and its meaning can be defined by the user. When all the routers in the domain have the common understanding what the particular algorithm value represents, the computation for such algorithm is consistent and the traffic is not subject to looping. Here, since the meaning of the algorithm is not defined by any standard, but is defined by the user, it is called as Flexible Algorithm.

## Flexible Algorithm Support Advertisement

An algorithm defines how the best path is computed by IGP. Routers advertise the support for the algorithm as a node capability. Prefix-SIDs are also advertised with an algorithm value and are tightly coupled with the algorithm itself.

An algorithm is a one octet value. Values from 128 to 255 are reserved for user defined values and are used for Flexible Algorithm representation.

## Flexible Algorithm Definition Advertisement

To guarantee the loop free forwarding for paths computed for a particular Flexible Algorithm, all routers in the network must share the same definition of the Flexible Algorithm. This is achieved by dedicated router(s) advertising the definition of each Flexible Algorithm. Such advertisement is associated with the priority to make sure that all routers will agree on a single and consistent definition for each Flexible Algorithm.

Definition of Flexible Algorithm includes:

- Metric type
- Affinity constraints

To enable the router to advertise the definition for the particular Flexible Algorithm, **advertise-definition** command is used. At least one router in the area, preferably two for redundancy, must advertise the Flexible Algorithm definition. Without the valid definition being advertised, the Flexible Algorithm will not be functional.

## Flexible Algorithm Prefix-SID Advertisement

To forward traffic on a Flexible Algorithm specific path, all routers participating in the Flexible Algorithm install a MPLS labeled path for the Flexible Algorithm specific prefix-SID. This Flexible Algorithm specific prefix-SID is advertised for the prefix. Only prefixes for which the Flexible Algorithm specific Prefix-SID is advertised, is subject to Flexible Algorithm specific forwarding.

### Inter-Area Leaking

Effective Cisco IOS XE Bengaluru 17.4.1, Flexible Algorithm SIDs and prefixes are leaked between IS-IS areas. However, only the prefixes that are reachable by Layer1 or Layer2 routers are leaked. Similarly, only SIDs that are reachable in a given Flexible Algorithm are leaked.

For example, consider a prefix P:

- that originated in Layer1 and leaked in to Layer2
- has SID value = 128 in Flexible Algorithm 128, and SID value = 129 in Flexible Algorithm 129
- for which Layer1 path exist only for SID value = 128, but not for SID value = 129

As a result of the above conditions, only SID 128 is leaked from Layer1 to Layer2 and not SID 129.

## Calculation of Flexible Algorithm Path

A router may compute path for multiple Flexible Algorithms. A router must be configured to support particular Flexible Algorithm before it can compute any path for such Flexible Algorithm. A router must have a valid definition of the Flexible Algorithm before such Flexible Algorithm is used.

When computing the shortest path tree for particular Flexible Algorithm:

- All nodes that do not advertise support for such Flexible Algorithm are pruned from the topology.
- If the Flexible Algorithm definition includes affinities that are excluded, then all links for which any of such affinities are advertised are pruned from the topology.
- Router uses the metric that is part of the Flexible Algorithm definition. If the metric is not advertised for the particular link, that link is pruned from the topology.

Effective Cisco IOS XE Bengaluru 17.4.1, for IS-IS, Loop Free Alternate (LFA) paths, TI-LFA backup paths, and Microloop Avoidance paths for a Flexible Algorithm are computed using the same constraints as the calculation of the primary paths for such Flexible Algorithm. These paths use Prefix-SIDs advertised specifically for such Flexible Algorithm to enforce a backup or microloop avoidance path.

## OSPF Flexible Algorithm TI-LFA

Table 13: Feature History

Feature Name	Release Information	Feature Description
OSPF Flexible Algorithm (Ph2): Topology-Independent Loop-Free Alternate (TI-LFA) Path	Cisco IOS XE Bengaluru 17.6.1	This feature allows you to configure the Loop-Free Alternate (LFA) and TI-LFA backup or repair paths for a Flexible Algorithm. The backup path is computed based on the constraints and metrics of the primary path. Prior to Cisco IOS XE Bengaluru 17.6.1, OSPF Flexible Algorithm supported only the primary path.

Effective Cisco IOS XE Bengaluru 17.6.1, for OSPFv2, OSPF Flexible Algorithm supports Loop Free Alternate (LFA) paths and Topology-Independent Loop-Free Alternate (TI-LFA) backup or repair paths. The constraints and metrics of the primary path applies for the computation of the backup path.

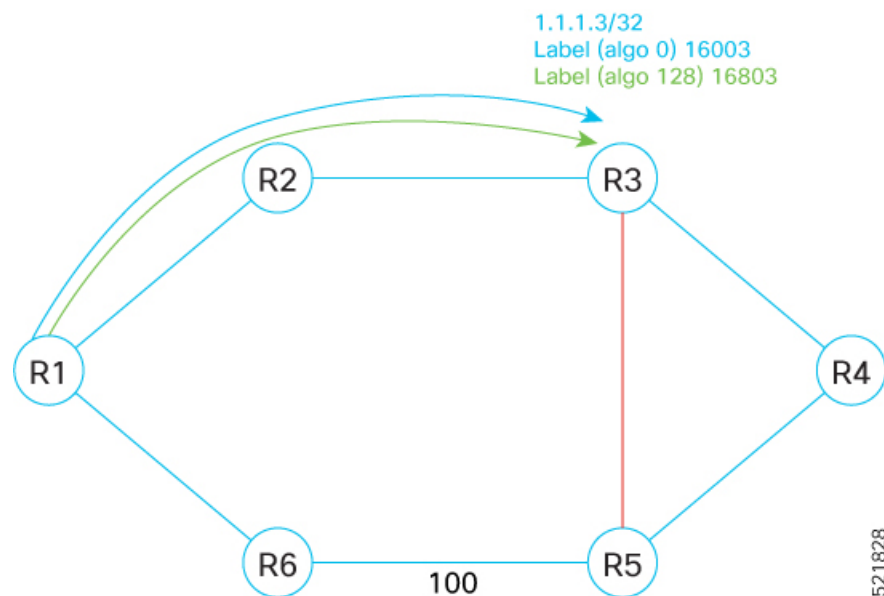
### OSPF Flexible Algorithm TI-LFA Computation

- The TI-LFA algorithm applies to the topology that you have included in the Flexible Algorithm.
- The backup path is expressed with the prefix-SIDs and optimized based on the Flexible Algorithm.

#### Examples

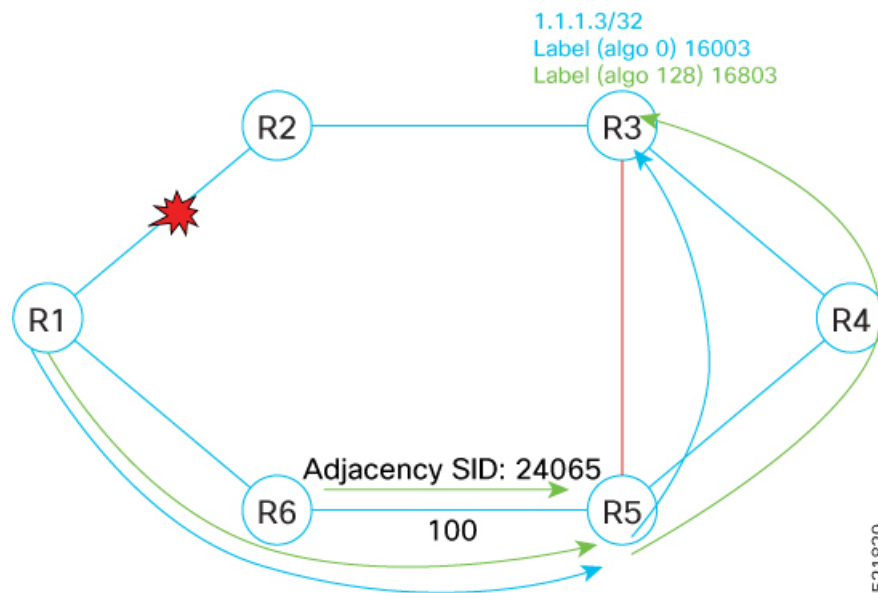
The following examples show how the Flexible Algorithm calculates the primary and TI-LFA backup paths:

Figure 23: Primary Path Per Flexible Algorithm



For the prefix 1.1.1.3/32, Algo 0 is assigned the prefix-SID 16003 and Algo 128 is assigned the prefix-SID 16803. Each prefix-SID is related to an algorithm. If a node advertises participation in a Flexible Algorithm, it also advertises a prefix-SID. All routers participate in Algo 128, which is defined as minimum IGP metric and avoid red affinity. The default IGP link metric is 1:10. The primary path for R1 to R3 is through R2.

Figure 24: TI-LFA Backup Path Per Flexible Algorithm



Assume that the link between R1 and R2 is down. The label 24065 is the adjacency SID for the link between R6 and R5. The backup path for Algo 0 is <24065, 16003> through R6.

The backup path for Algo 128 is <24065, 16803> through R6. Therefore, the usage of Algo 128 prefix-SID 16803 ensures that the Algo 128 backup path also avoids the red affinity and uses R1→R6→R5→R4→R3.

## Configuring OSPF Flexible Algorithm TI-LFA

**Step 1** Configure the Flexible Algorithm TI-LFA functionality.

```
Router(config)#router ospf 1
Router(config-router)#router-id 255.255.255.255
Router(config-router)#nsr
Router(config-router)#segment-routing area 0 mpls
Router(config-router)#segment-routing mpls
Router(config-router)#segment-routing prefix-sid-map advertise-local
Router(config-router)#fast-reroute per-prefix enable prefix-priority low
Router(config-router)#fast-reroute per-prefix ti-lfa
Router(config-router)#flex-algo 129
Router(config-ospf-fa)#fast-reroute disable
Router(config-ospf-fa)#metric-type te-metric
Router(config-ospf-fa)#affinity include-all
Router(config-ospf-fa-aff)#name GREEN
Router(config-ospf-fa-aff)#name RED
```

```
Router(config-ospf-fa-aff)#affinity-map
Router(config-ospf-aff-map)#name GREEN bit-position 130
Router(config-ospf-aff-map)#name RED bit-position 129
Router(config-ospf-aff-map)#end
```

**Note** You can use the **fast-reroute disable** command to disable TI-LFA for a particular Flexible Algorithm.

**Step 2** Attach the OSPF Flex-Algo to the SR-TE policy to steer traffic.

```
Router(config)#segment-routing traffic-eng
Router(config-srte)#policy FLEXALGO1
Router(config-srte-policy)#color 1 end-point 6.6.6.6
Router(config-srte-policy)#candidate-paths
Router(config-srte-policy-path)#preference 1
Router(config-srte-policy-path-pref)#constraints
Router(config-srte-policy-path-pref-constr)#segments
Router(config-srte-policy-path-pref-constr-seg)#dataplane mpls
Router(config-srte-policy-path-pref-constr-seg)#algorithm 129
Router(config-srte-policy-path-pref-constr-seg)#!
Router(config-srte-policy-path-pref-constr-seg)#!
Router(config-srte-policy-path-pref-constr-seg)#dynamic
Router(config-srte-policy-path-pref-dyn)#end
```

## Verifying the OSPF Flexible Algorithm TI-LFA

Verify the OSPF Flexible Algorithm TI-LFA:

**Step 1** Use the **show segment-routing traffic-eng policy name** command to verify the IP address of the endpoint, the algorithm attached, tunnel ID, and the event history.

```
Router#show segment-routing traffic-eng policy name <FLEXALGO1> [detail]
```

```
Name: FLEXALGO1 (Color: 1 End-point: 6.6.6.6)
  Owners : CLI
  Status:
    Admin: up, Operational: up for 00:00:04 (since 05-07 15:26:18.320)
  Candidate-paths:
    Preference 1 (CLI):
      Constraints:
        Algorithm: 129--- Algo 129 is attached
      Dynamic (active)
        17666 [Prefix-SID, 6.6.6.6]
  Attributes:
    Binding SID: 16
    Allocation mode: dynamic
    State: Programmed
  Tunnel ID: 65536 (Interface Handle: 0x21)
  Per owner configs:
    CLI
    Binding SID: dynamic
  Stats:
    5 minute output rate 0 bits/sec, 0 packets/sec
    Packets: 0 Bytes: 0

  Event history:
```

Timestamp	Client	Event type	Context: Value
05-05 11:50:51.019	CLI	Policy created	Name: CLI
05-05 11:50:51.089	CLI	Set colour	Colour: 1

```

05-07 15:25:35.625      CLI      Set colour      Colour: 1
05-07 15:25:35.625      CLI      Set end point   End-point: 6.6.6.6
05-07 15:25:35.640      FH Resolution  Policy state UP   Status: PATH RESOLVED
CP: 1
05-07 15:25:36.303      FH Resolution  REOPT triggered   Status: REOPTIMIZED
CP: 1
05-07 15:26:18.321      CLI      Policy ADMIN DOWN shutdown: FLEXALGO1
05-07 15:26:19.173      CLI      Policy state DOWN no shutdown: FLEXALGO1
05-07 15:26:19.177      FH Resolution  Policy state UP   Status: PATH RESOLVED
CP: 1

```

**Step 2** Use the **show ip ospf topology route algorithm** command to display the path information of the route computed based on route calculation. The command displays details such as the attached label, the algorithm ID, SID, total metric, primary path (label), and the backup path (label and TI-LFA tunnel).

```

Router#show ip ospf topology route algorithm 129 6.6.6.6/32 detail
Route Table of OSPF-1 with router ID 255.255.255.255 (VRF global)
Algorithm 129

Route entry for 6.6.6.6/32, Metric 102, SID 666, Label 17666
Priority : Medium
Flags : New

Route type : Intra
Flags: Inuse
Prefix Contrib Algo 129 SID 666
Total Metric : 102
Contrib Flags : Inuse, Reachable
SID Flags : Exp Null, PHP off, Index, Global, Node, Valid
Path: 102.0.0.1, from 4.4.4.4, via GigabitEthernet0/5/1----- primary path
Out Label : 17666
Weight : 0

Backup path: TI-LFA, Repair-List: P node: 5.5.5.5 Label: 17555
Backup Tunnel: MPLS-SR-Tunnel7----- TILFA Tunnel
101.0.0.1, from 6.6.6.6, via BDI1001,
Out Label: 17666
Attributes: Metric: 203, LC Disjoint, Node Protect, Interface Disjoint, SRLG Disjoint

Path: 122.0.0.1, from 4.4.4.4, via BDI1022
Out Label : 17666
Weight : 0

Backup path: TI-LFA, Repair-List: P node: 5.5.5.5 Label: 17555
Backup Tunnel: MPLS-SR-Tunnel7
101.0.0.1, from 6.6.6.6, via BDI1001,
Out Label: 17666
Attributes: Metric: 203, Node Protect, Interface Disjoint, SRLG Disjoint

```

**Step 3** Use the **show ip ospf topology fast-reroute tunnel** command to display details such as the flex-algo TI-LFA tunnels, interface, next hop, mid/end points, and label.

```

Router#show ip ospf topology fast-reroute tunnel

OSPF Router with ID (255.255.255.255) (Process ID 1)

Tunnel      Interface      Next Hop      Mid/End Point  Label
-----

```

MPLS-SR-Tunnel7	BD1001	101.0.0.1	5.5.5.5	17555
MPLS-SR-Tunnel8	BD1001	101.0.0.1	7.7.7.7	17777

**Step 4** Use the **show mpls infrastructure lfd lte** <flex-algo label> command to view the next hop outgoing interface/IP address.

```
RSP2-12#show mpls infrastructure lfd lte 17111
17111 [table 0]
  ldm: LTE Broker, OSPF
  flags: nsf | ver2 | multi-bind | srlbls (0x1B)
  plist: NONIP (0x6A3B5FB8) 0-11.11.11.11/32-33028
  output chain:
    label 17111-(local:17111)
    TAG adj out of GigabitEthernet0/5/1, addr 102.0.0.1 7F083B00
```

**Step 5** Use the **show ip ospf topology summary** command to display information such as number of areas, algos, nodes, the algo label, and the metric type.

```
show ip ospf [process-id] [area-id] topology summary [algorithm algorithm-number]
```

```
Router#show ip ospf topology summary
      Process OSPF-1

Instance global
Router ID       : 255.255.255.255
Number of Areas : 1
Number of Algos : 1
Max Path count  : 4
Route count     : 10
SR Global Block : 17000 - 23999
Area 0
  Number of Nodes : 11
  Algo 129
    FAD Advertising Router : 199.199.199.199
    FAD Area ID : 0
    Algo Type   : 0
    Metric Type : 1
    Include Any Affinity : 0x00000000 00000000 00000000 00000000 00000000
```

## Installation of Forwarding Entries for Flexible Algorithm Paths

Flexible Algorithm path to any prefix must be installed in the forwarding using the Prefix-SID that was advertised for such Flexible Algorithm. If the Prefix-SID for Flexible Algorithm is not known, such Flexible Algorithm path is not installed in forwarding for such prefix..

Only MPLS to MPLS entries are installed for a Flexible Algorithm path. No IP to IP or IP to MPLS entries are installed. These follow the native IPG paths computed based on the default algorithm and regular IGP metrics.

Effective Cisco IOS XE Bengaluru 17.4.1, you can selectively filter the paths that are installed to the MFI by using the configuration command **distribute-list filter name in**. See [Configuring Selective Path Filtering](#) for configuration example. This feature is only supported for IS-IS Flexible Algorithm.

## Flexible Algorithm Configurations

This section describes various configurations that are required to support the SR Flexible Algorithm functionality.

The following IS-IS and OSPF configuration sub-mode is used to configure Flexible Algorithm:

```
flex-algo algorithm number
```

*algorithm number* —value from 128 to 255

### Commands under Flexible Algorithm Configuration Mode

The following commands are used to configure Flexible Algorithm definition under the flex-algo sub-mode:

- To set the metric delay:

- In IS-IS

```
metric-type delay
```




---

**Note** By default the regular IGP metric is used. If delay metric is enabled, the advertised delay on the link is used as a metric for Flexible Algorithm computation.

---

- In OSPF

```
metric-type {delay | te-metric | igp-metric}
```

- To set the affinity:

- In IS-IS

```
affinity
exclude-any
```

```
name affinity-name
```

- In OSPF

```
affinity {exclude-any | include-any | include-all}
name affinity-name
```

*affinity-name*—name of the affinity-map

- To set the priority in both IS-IS and OSPF:

```
priority priority value
```

*priority value*—priority used during the Flexible Algorithm definition election.

The following command is used to enable advertisement of the Flexible Algorithm definition in IS-IS and OSPF:

```
advertise-definition
```

### Commands for Affinity Configuration

The following commands are used for defining the affinity-map. Affinity-map associates the name with the particular bit positions in the Extended Admin Group bitmask.

- In IS-IS:



```
affinity-map affinity name bit-position bit number
```

- In OSPF:

```
affinity-map
name affinity name bit-position bit number
```

- *affinity name*—name of the affinity-map.
- *bit number*—bit position in the Extended Admin Group bitmask.

The following command is used to associate the affinity with an interface:

- IS-IS

```
isis affinity flex-algo
name affinity-name
```

- OSPF

```
ip ospf affinity flex-algo
name affinity-name

affinity-name—name of the affinity-map
```

### Command for Prefix-SID for Flexible Algorithm Configuration

The following command is used to advertise prefix-SID for flexible algorithm in IS-IS and OSPF:

```
segment-routing mpls
connected-prefix-sid-map
address-family ipv4 [ algorithm algorithm-number | strict-spf]
prefix/sid [index | absolute] sid value range range value
```

## Configuring IS-IS Flexible Algorithm

The following is an example of how to configure the IS-IS flexible algorithm.

```
router isis 1
net 49.0002.0000.0001.00
is-type level-1
metric-style wide
log-adjacency-changes
nsf cisco
distribute link-state
segment-routing mpls
segment-routing prefix-sid-map advertise-local
affinity-map blue bit-position 8
affinity-map green bit-position 201
affinity-map red bit-position 65

fast-reroute per-prefix level-1 all
fast-reroute tie-break level-1 node-protecting 100
fast-reroute tie-break level-1 srlg-disjoint 50
fast-reroute ti-lfa level-1
fast-reroute ti-lfa level-2
microloop avoidance segment-routing
microloop avoidance rib-update-delay 10000
```

```

flex-algo 129
  advertise-definition
  metric-type delay
  priority 120
  affinity
    exclude-any
      name red
  !
bfd all-interfaces
mpls traffic-eng router-id Loopback1
mpls traffic-eng level-1
mpls traffic-eng level-2

```

Use the **fast-reroute disable** command to disable TI-LFA.

## Configuring OSPF Flexible Algorithm

The following is an example of how to configure OSPF flexible algorithm.

```

interface Gi0/0/0
  ip ospf affinity flex-algo
  name red
  !
interface Ethernet0/1
  ip ospf affinity flex-algo
  name blue
  name red
  !
interface Ethernet0/2
  ip ospf affinity flex-algo
  name blue
  !

router ospf 1

  affinity-map
    name red bit-position 65
    name blue bit-position 8
    name green bit-position 201
  !
  flex-algo 128
    affinity exclude-any
      name red
      name blue
  !
  affinity include-any
    name green
  !
  !
  flex-algo 129
    affinity exclude-any
      name green
  !
  affinity include-all
    name red
    name blue
  !
  advertise-definition
  priority 20
  metric-type delay
  !

```

## Configuring Prefix-SID

```
segment-routing mpls
set-attributes
  address-family ipv4
  sr-label-preferred
  exit-address-family
  global-block 17000 23999
  connected-prefix-sid-map
  address-family ipv4
  1.1.1.1/32 index 18 range 1
  exit-address-family
  address-family ipv4 algorithm 129
  1.1.1.1/32 index 38 range 1
  exit-address-family
```

## Configuring SRTE-ODN Association

```
segment-routing traffic-eng
on-demand color 100
authorize
candidate-paths
preference 100
constraints
  segments
    dataplane mpls
    algorithm 129
  !
!
dynamic
  metric
    type delay
  !
!
```

## Configuring the Interface for Flexible Algorithm

```
interface GigabitEthernet0/0/6
ip address 11.11.11.1 255.255.255.0
ip router isis 1
mpls ip
mpls traffic-eng tunnels
bfd template pw_bfd
isis network point-to-point
isis affinity flex-algo
name red
!
```

## Configuring BGP

```
router bgp 100
  bgp router-id 1.1.1.1
  bgp log-neighbor-changes
  bgp graceful-restart
  neighbor 2.2.2.2 remote-as 100
  neighbor 2.2.2.2 ha-mode sso
  neighbor 2.2.2.2 update-source Loopback1
  !
```

```

address-family ipv4
  neighbor 2.2.2.2 activate
exit-address-family
!
address-family vpnv4
  neighbor 2.2.2.2 activate
  neighbor 2.2.2.2 send-community both
  neighbor 2.2.2.2 route-map BGP_TE_MAP out
exit-address-family
!
address-family ipv4 vrf SR
  redistribute connected
  neighbor 132.1.1.1 remote-as 101
  neighbor 132.1.1.1 activate
exit-address-family
!

```

## Configuring Selective Path Filtering

The following is an example of how you can selectively filter the paths that are installed to the MFI.

```

R1-ASR900-2011(config)#show run int lo1
Building configuration...

```

```

Current configuration : 95 bytes
!
interface Loopback1
 ip address 1.1.1.1 255.255.255.255
 ip router isis 1
 isis tag 111
end

```

```

On R6-node:
+++++
!
route-map block deny 10
 match tag 111
!
route-map block permit 100
!

router isis 1
!
 flex-algo 135
  advertise-definition
  metric-type delay
  priority 120
  affinity
  exclude-any
  name green
!
!
distribute-list route-map block in
 bfd all-interfaces
 mpls traffic-eng router-id Loopback1
 mpls traffic-eng level-1
 mpls traffic-eng level-2
!
R6-RSP3-2038(config)#show isis rib flex-algo 129

IPv4 local RIB for IS-IS process 1

```

## Verifying the Flexible Algorithm Configuration

```
show isis flex-algo 129
Tag 1:
IS-IS Flex-Algo Database
  Flex-Algo count: 7

Flex-Algo 129:
  IS-IS Level-1
    Definition Priority: 222
    Definition Source: R2-RSP3-2015.00, (Local)
    Definition Equal to Local: Yes
    Definition Metric Type: Delay
```

```

Definition Flex-Algo Prefix Metric: No
Disabled: No
Microloop Avoidance Timer Running: No
Local Priority: 222
FRR Disabled: No
Microloop Avoidance Disabled: No

```

Use the command **show isis rib flex-algo** *value* to view all the IS-IS local RIB information.

```

show isis rib flex-algo 129
IPv4 local RIB for IS-IS process 1

IPV4 unicast topology base (TID 0, TOPOID 0x0) =====
Repair path attributes:
  DS - Downstream, LC - Linecard-Disjoint, NP - Node-Protecting
  PP - Primary-Path, SR - SRLG-Disjoint

Flex-algo 129

1.1.1.1/32 prefix attr X:0 R:0 N:1 source router id: 1.1.1.1 SID index 38 - Bound
[115/L1/113] via 11.11.11.1(GigabitEthernet0/4/6) R1-ASR920-2011.00-00, from 1.1.1.1, tag
0
  LSP 6/6/351(351), prefix attr: X:0 R:0 N:1
  Source router id: 1.1.1.1
  Prefix-SID index: 38, R:0 N:1 P:0 E:0 V:0 L:0
  label: implicit-null
  repair path: 20.20.20.2 (GigabitEthernet0/4/7) metric: 117 (DS,SR)
    local LFA
    label: implicit-null
    repair source: R1-ASR920-2011, LSP 6

2.2.2.2/32 prefix attr X:0 R:0 N:1 source router id: 2.2.2.2 SID index 39 - Bound
[115/L1/24] via 13.13.13.2(GigabitEthernet0/1/5) R4-RSP3-2036.00-00, from 2.2.2.2, tag 0
  LSP 2/3/345(345), prefix attr: X:0 R:0 N:1
  Source router id: 2.2.2.2
  Prefix-SID index: 39, R:0 N:1 P:0 E:0 V:0 L:0
  label: 17039
  repair path: 4.4.4.4 (MPLS-SR-Tunnel4) metric: 170 (DS,NP,SR)
    next-hop: 20.20.20.2 (GigabitEthernet0/4/7)
    TI-LFA node/SRLG-protecting, SRLG-protecting
    SRGB: 17000, range: 7000 prefix-SID index: 39, R:0 N:1 P:0 E:0 V:0 L:0
    label: 17039
    P node: R3-RSP2-2013[4.4.4.4], label: 17221
    repair source: R6-RSP3-2038, LSP 3

4.4.4.4/32 prefix attr X:0 R:0 N:1 source router id: 4.4.4.4 SID index 221 - Bound
[115/L1/172] via 13.13.13.2(GigabitEthernet0/1/5) R4-RSP3-2036.00-00, from 4.4.4.4, tag
0
  LSP 2/7/24(24), prefix attr: X:0 R:0 N:1
  Source router id: 4.4.4.4
  Prefix-SID index: 221, R:0 N:1 P:0 E:0 V:0 L:0
  label: 17221
  repair path: 20.20.20.2 (GigabitEthernet0/4/7) metric: 184 (DS,NP,SR)
    local LFA
    label: 17221
    repair source: R3-RSP2-2013, LSP 7

5.5.5.5/32 prefix attr X:0 R:0 N:1 source router id: 5.5.5.5 SID index 222 - Bound
[115/L1/17] via 13.13.13.2(GigabitEthernet0/1/5) R4-RSP3-2036.00-00, from 5.5.5.5, tag 0
  LSP 2/2/347(347), prefix attr: X:0 R:0 N:1
  Source router id: 5.5.5.5
  Prefix-SID index: 222, R:0 N:1 P:0 E:0 V:0 L:0
  label: implicit-null
  repair path: 4.4.4.4 (MPLS-SR-Tunnel4) metric: 170 (DS,SR)

```

```

next-hop: 20.20.20.2 (GigabitEthernet0/4/7)
TI-LFA SRLG-protecting
SRGB: 17000, range: 7000 prefix-SID index: 222, R:0 N:1 P:0 E:0 V:0 L:0
label: 17222
P node: R3-RSP2-2013[4.4.4.4], label: 17221
repair source: R4-RSP3-2036, LSP 2

6.6.6.6/32 prefix attr X:0 R:0 N:1 source router id: 6.6.6.6 SID index 333 - Bound
[115/L1/122] via 13.13.13.2(GigabitEthernet0/1/5) R4-RSP3-2036.00-00, from 6.6.6.6, tag
0
LSP 2/4/351(351), prefix attr: X:0 R:0 N:1
Source router id: 6.6.6.6
Prefix-SID index: 333, R:0 N:1 P:0 E:0 V:0 L:0
label: 17333
repair path: 4.4.4.4 (MPLS-SR-Tunnel4) metric: 170 (DS,NP,SR)
next-hop: 20.20.20.2 (GigabitEthernet0/4/7)
TI-LFA node/SRLG-protecting, SRLG-protecting
SRGB: 17000, range: 7000 prefix-SID index: 333, R:0 N:1 P:0 E:0 V:0 L:0
label: 17333
P node: R3-RSP2-2013[4.4.4.4], label: 17221
repair source: R5-ASR920-2012, LSP 4

```

Use the command **show isis topo flex-algo** *value* for information regarding the IS-IS paths to intermediate systems.

```

show isis topo flex-algo 129
Tag 1:
IS-IS TID 0 paths to level-1 routers
Flex-algo 129
System Id          Metric      Next-Hop          Interface        SNPA
920_1              3          RSP2_2           Gi0/15/0        e8ed.f3b8.f804
RSP3_R1            **
RSP2_1             2          RSP2_2           Gi0/15/0        e8ed.f3b8.f804
RSP3_R2            **
RSP2_2             1          RSP2_2           Gi0/15/0        e8ed.f3b8.f804
RSP3_R3            --

```

Use the command **show isis fast-reroute ti-lfa tunnel** for information regarding the IS-IS TI-LFA tunnels.

```

show isis fast-reroute ti-lfa tunnel
Tag null:
Fast-Reroute TI-LFA Tunnels:
Tunnel Interface Next Hop      End Point      Label      End Point Host
Tag 1:
Fast-Reroute TI-LFA Tunnels:

Tunnel Interface Next Hop      End Point      Label      End Point Host
MP2      Gi0/0/6      12.12.12.2      2.2.2.2        17019      RSP3_R3
MP5      Gi0/0/5      11.11.11.2      2.2.2.2        17019      RSP3_R3
MP3      Gi0/0/6      12.12.12.2      6.6.6.6        17333      RSP2_2
                2.2.2.2        16         RSP3_R3
MP9      Gi0/0/5      11.11.11.2      2.2.2.2        17039      RSP3_R3
MP1      Gi0/0/6      12.12.12.2      6.6.6.6        20333      RSP2_2
                2.2.2.2        16         RSP3_R3
MP6      Gi0/0/5      11.11.11.2      2.2.2.2        17049      RSP3_R3

```

Use the command **show isis flex-algo** *[flex-algorithm-id]* to display the configured flexible algorithm and the defined parameters.

```
R1#show isis flex-algo 255
Tag Ring#1:
IS-IS Flex-Algo Database
Flex-Algo count: 1

Flex-Algo 255:
  IS-IS Level-1
    Definition Priority: 11
    Definition Source: PE23.00
    Definition Equal to Local: Yes
    Definition Metric Type: IGP
    Definition Include-any Affinity:
      0x00000000 0x00000000 0x00000002
    Definition Flex-Algo Prefix Metric: No
    Disabled: No
    Microloop Avoidance Timer Running: No
  IS-IS Level-2
    Definition Priority: 11
    Definition Source: PE23.00
    Definition Equal to Local: Yes
    Definition Metric Type: IGP
    Definition Include-any Affinity:
      0x00000000 0x00000000 0x00000002
    Definition Flex-Algo Prefix Metric: No
    Disabled: No
    Microloop Avoidance Timer Running: No
  Local Priority: 11
  FRR Disabled: No
  Microloop Avoidance Disabled: No
```

Use the command **show ip ospf topology summary internal** to display instance level and area level information.

**show ip ospf** *[process-id] [area-id] topology summary [algorithmalgorithm-number]*  
**[internal]**

```
R1#show ip ospf topology summary internal
Process OSPF-10

Instance global
Router ID       : 1.1.1.1
Number of Areas : 2
Number of Algos : 2
Max Path count  : 4
Route count     : 4
Algo Status
  Algo 128 : In-use
  Algo 129 : In-use
Route Work Queue Sizes
  Critical : 0, High      : 0
  Medium   : 0, Low       : 0
SR Global Block : 16000 - 24000
Area
  Number of Nodes : 8
Algo 128
  FAD Advertising Router : 1.1.1.1
  FAD Area ID : 0
  Algo Type   : 0
  Metric Type : 0
  Include Any Affinity : 0x00000002
  SPF count   : 10
```



```

SPF Trigger Flags : Valid, Explicit FAD
Algo 129
FAD Advertising Router : 1.1.1.1
FAD Area ID : 0
  Algo Type      : 0
  Metric Type    : 0
  Exclude Affinity : 0x00000002
SPF count : 11
SPF Trigger Flags : Valid, Explicit FAD
Area
Number of Nodes : 2
Algo 128
FAD Advertising Router : 1.1.1.1
FAD Area ID : 1
  Algo Type      : 0
  Metric Type    : 0
  Include Any Affinity : 0x00000002
SPF count : 4
SPF Trigger Flags : Valid, Explicit FAD
Algo 129
FAD Advertising Router : 1.1.1.1
FAD Area ID : 1
  Algo Type      : 0
  Metric Type    : 0
  Exclude Affinity : 0x00000002
SPF count : 4
SPF Trigger Flags : Valid, Explicit FAD

```

Use the command **show ip ospf topology** to display the node and link information compiled from the LSAs.

**show ip ospf** [*process-id*] [*area-id*] **topology** [*algorithm* *algorithm-number*] [*node-id* *router-id*] [*link-address* *link-address*] [**detail**]

```

R1#show ip ospf topology
    Process OSPF-10

Instance : global
Router ID : 1.1.1.1
Area : (8 nodes)
  Node : 1.2.0.2 (pseudo) (2 links)
    Link : 1.1.1.1 0.0.0.0 Transit
    Link : 1.1.1.2 0.0.0.0 Transit
  Node : 1.1.1.1 (root) (3 links) ABR
    Algos supported: 128, 129
    Flex Algo Definition: 128
    Flex Algo Definition: 129
    Link : 1.1.1.6 0.0.0.2 Point-to-point
    Link : 1.1.1.6 1.6.1.1 Point-to-point
    Link : 1.2.0.2 1.2.0.1 Transit
  Node : 1.1.1.2 (3 links)
    Algos supported: 128
    Link : 1.1.1.3 2.3.0.2 Point-to-point
    Link : 1.1.1.54 2.5.0.2 Point-to-point
    Link : 1.2.0.2 1.2.0.2 Transit
  Node : 1.1.1.3 (2 links)
    Algos supported: 128
    Link : 1.1.1.2 2.3.0.3 Point-to-point
    Link : 1.1.1.4 3.4.0.3 Point-to-point
  Node : 1.1.1.4 (3 links) ABR, ASBR
    Algos supported: 128, 129
    Link : 1.1.1.3 3.4.0.4 Point-to-point
    Link : 1.1.1.9 0.0.0.3 Point-to-point
    Link : 1.1.1.54 4.5.0.4 Point-to-point
  Node : 1.1.1.6 (4 links)
    Algos supported: 129

```

```

Link : 1.1.1.1 0.0.0.2 Point-to-point
Link : 1.1.1.1 1.6.1.6 Point-to-point
Link : 1.1.1.54 5.6.0.6 Point-to-point
Link : 1.1.1.54 5.6.1.6 Point-to-point
Node : 1.1.1.9 (1 links) ABR
Link : 1.1.1.4 0.0.0.3 Point-to-point
Node : 1.1.1.54 (4 links)
Algos supported: 129
Link : 1.1.1.2 2.5.0.5 Point-to-point
Link : 1.1.1.4 4.5.0.5 Point-to-point
Link : 1.1.1.6 5.6.0.5 Point-to-point
Link : 1.1.1.6 5.6.1.5 Point-to-point
Area : (2 nodes)
Node : 1.1.1.1 (root) (1 links) ABR
Algos supported: 128, 129
Flex Algo Definition: 128
Flex Algo Definition: 129
Link : 1.1.1.8 1.8.0.1 Point-to-point
Node : 1.1.1.8 (1 links) ASBR
Link : 1.1.1.1 1.8.0.8 Point-to-point

R1#show ip ospf topo node-id 1.1.1.1 de
Process OSPF-10

Instance : global
Router ID : 1.1.1.1
Area : (8 nodes)
Node : 1.1.1.1 (root) (3 links) ABR
No. of algos : 2
Algos supported: 128, 129
SRGB List:
16000 - 24000
Oper Flags : 0x00
Contributing LSAs : Router, Router Info

Flex Algo Definition: 128
Algo Type : 0
Metric Type : 0
Priority : 10
Include Any Affinity : 0x00000002
Oper Flags : 0x00

Flex Algo Definition: 129
Algo Type : 0
Metric Type : 0
Priority : 128
Exclude Affinity : 0x00000002
Oper Flags : 0x00

Link : 1.1.1.6 0.0.0.2 Point-to-point
Metric : 10
Remote link: 1.1.1.1 0.0.0.2 Point-to-point
Oper Flags : 0x00
Link Attributes:
LSID : 8.0.0.2
Local-Remote Interface ID : (2, 2)
SR TE Metric : 40
Contributing LSAs : Router, Ext. Link

Link : 1.1.1.6 1.6.1.1 Point-to-point
Metric : 10
Subnet mask: 255.255.255.0
Remote link: 1.1.1.1 1.6.1.6 Point-to-point
Oper Flags : 0x00

```

```

Link Attributes:
  LSID      : 8.0.0.5
  Local-Remote Interface ID : (5, 5)
  Remote IPv4 Address: 1.6.1.6
  SR TE Metric : 40
  Number of SRLGs : (2)
    [1]: 1          [2]: 6
  Contributing LSAs : Router, Ext. Link

Link : 1.2.0.2 1.2.0.1 Transit
  Metric      : 10
  Remote link: 1.1.1.1 0.0.0.0 Transit
  Oper Flags : 0x00
Link Attributes:
  LSID      : 8.0.0.3
  SR TE Metric : 40
  Ext. Admin Group: 0x00000002
  Contributing LSAs : Router, Ext. Link

Area : (2 nodes)
Node : 1.1.1.1 (root) (1 links) ABR
No. of algos : 2
  Algos supported: 128, 129
  SRGB List:
    16000 - 24000
  Oper Flags : 0x00
  Contributing LSAs : Router, Router Info

Flex Algo Definition: 128
  Algo Type : 0
  Metric Type : 0
  Priority : 10
  Include Any Affinity : 0x00000002
  Oper Flags : 0x00

Flex Algo Definition: 129
  Algo Type : 0
  Metric Type : 0
  Priority : 128
  Exclude Affinity : 0x00000002
  Oper Flags : 0x00

Link : 1.1.1.8 1.8.0.1 Point-to-point
  Metric      : 10
  Subnet mask: 255.255.255.0
  Remote link: 1.1.1.1 1.8.0.8 Point-to-point
  Oper Flags : 0x00
Link Attributes:
  LSID      : 8.0.0.4
  Local-Remote Interface ID : (4, 4)
  Remote IPv4 Address: 1.8.0.8
  SR TE Metric : 20
  Min Delay : 499
  Contributing LSAs : Router, Ext. Link

```

Use the command **show ip ospf topology prefix** display the node and prefix information compiled from the LSAs.

```
show ip ospf [process-id] [area-id] topology prefix [algorithmalgorithm-number]
[node-id router-id] [prefix/mask] [detail]
```

```
R1#show ip ospf topology prefix
Process OSPF-10
```

```

Instance : global
Router ID : 1.1.1.1
Area : (8 nodes)
  Node : 1.2.0.2 (pseudo) (2 links)
  Node : 1.1.1.1 (root) (3 links) ABR
    Algos supported: 128, 129
    Flex Algo Definition: 128
    Flex Algo Definition: 129
  Node : 1.1.1.2 (3 links)
    Algos supported: 128
  Node : 1.1.1.3 (2 links)
    Algos supported: 128
    Prefix : 1.1.1.34/32
  Node : 1.1.1.4 (3 links) ABR, ASBR
    Algos supported: 128, 129
    Prefix : 1.1.1.4/32
    Prefix : 1.1.1.34/32
    Prefix : 1.1.1.45/32
  Node : 1.1.1.6 (4 links)
    Algos supported: 129
  Node : 1.1.1.9 (1 links) ABR
  Node : 1.1.1.54 (4 links)
    Algos supported: 129
    Prefix : 1.1.1.54/32
Area : (2 nodes)
  Node : 1.1.1.1 (root) (1 links) ABR
    Algos supported: 128, 129
    Flex Algo Definition: 128
    Flex Algo Definition: 129
  Node : 1.1.1.8 (1 links) ASBR

R1#show ip ospf topo prefix 1.1.1.4/32 de
  Process OSPF-10

Instance : global
Router ID : 1.1.1.1
Area : (8 nodes)
  Node : 1.2.0.2 (pseudo) (2 links)
    No. of algos : 0
    Oper Flags : 0x00
    Contributing LSAs : Network

  Node : 1.1.1.1 (root) (3 links) ABR
    No. of algos : 2
    Algos supported: 128, 129
    SRGB List:
      16000 - 24000
    Oper Flags : 0x00
    Contributing LSAs : Router, Router Info

    Flex Algo Definition: 128
      Algo Type : 0
      Metric Type : 0
      Priority : 10
      Include Any Affinity : 0x00000002
      Oper Flags : 0x00

    Flex Algo Definition: 129
      Algo Type : 0
      Metric Type : 0
      Priority : 128
      Exclude Affinity : 0x00000002
      Oper Flags : 0x00

```

```
Node : 1.1.1.2 (3 links)
  No. of algos : 1
    Algos supported: 128
  SRGB List:
    16000 - 24000
  Oper Flags : 0x00
  Contributing LSAs : Router, Router Info

Node : 1.1.1.3 (2 links)
  No. of algos : 1
    Algos supported: 128
  SRGB List:
    16000 - 24000
  Oper Flags : 0x00
  Contributing LSAs : Router, Router Info

Node : 1.1.1.4 (3 links) ABR, ASBR
  No. of algos : 2
    Algos supported: 128, 129
  SRGB List:
    16000 - 30001
  Oper Flags : 0x00
  Contributing LSAs : Router, Router Info

Prefix : 1.1.1.4/32
Metric : 1
Route Type : Intra

  Algo 128 SID 132
    Total Metric : 31
    Contrib Flags : Inuse, Reachable
    SID Flags : Exp Null, PHP off, Index, Global, Node, Valid

  Algo 129 SID 133
    Total Metric : 31
    Contrib Flags : Inuse, Reachable
    SID Flags : Exp Null, PHP off, Index, Global, Node, Valid

Node : 1.1.1.6 (4 links)
  No. of algos : 1
    Algos supported: 129
  SRGB List:
    16000 - 24000
  Oper Flags : 0x00
  Contributing LSAs : Router, Router Info

Node : 1.1.1.9 (1 links) ABR
  No. of algos : 0
  SRGB List:
    16000 - 24000
  Oper Flags : 0x00
  Contributing LSAs : Router, Router Info

Node : 1.1.1.54 (4 links)
  No. of algos : 1
    Algos supported: 129
  SRGB List:
    16000 - 24000
  Oper Flags : 0x00
  Contributing LSAs : Router, Router Info

Area : (2 nodes)
  Node : 1.1.1.1 (root) (1 links) ABR
    No. of algos : 2
```

```

    Algos supported: 128, 129
    SRGB List:
    16000 - 24000
    Oper Flags : 0x00
    Contributing LSAs : Router, Router Info

    Flex Algo Definition: 128
    Algo Type   : 0
    Metric Type : 0
    Priority    : 10
    Include Any Affinity : 0x00000002
    Oper Flags  : 0x00

    Flex Algo Definition: 129
    Algo Type   : 0
    Metric Type : 0
    Priority    : 128
    Exclude Affinity : 0x00000002
    Oper Flags  : 0x00

    Node : 1.1.1.8 (1 links) ASBR
    No. of algos : 0
    SRGB List:
    16000 - 24000
    Oper Flags : 0x00
    Contributing LSAs : Router, Router Info

```

Use the command **show ip ospf topology path** display the path information of nodes computed by SPF.

**show ip ospf** [*process-id*] [*area-id*] **topology path** [*algorithm* *algorithm-number*]  
 [*node-id* *router-id*] [**detail**]

```

R1#show ip ospf topology path
    Process OSPF-10

    Instance global
    Router ID : 1.1.1.1
    Area
    Node 1.1.1.2
    Algo 128, Distance 10
    1.2.0.2, from 1.1.1.2, via Ethernet0/1
    Node 1.1.1.3
    Algo 128, Distance 20
    1.2.0.2, from 1.1.1.2, via Ethernet0/1
    Node 1.1.1.4 ABR, ASBR
    Algo 128, Distance 30
    1.2.0.2, from 1.1.1.2, via Ethernet0/1
    Algo 129, Distance 30
    1.1.1.6, from 1.1.1.6, via Ethernet0/0
    1.6.1.6, from 1.1.1.6, via Ethernet0/3
    Node 1.1.1.6
    Algo 129, Distance 10
    1.1.1.6, from 1.1.1.6, via Ethernet0/0
    1.6.1.6, from 1.1.1.6, via Ethernet0/3
    Node 1.1.1.54
    Algo 129, Distance 20
    1.1.1.6, from 1.1.1.6, via Ethernet0/0
    1.6.1.6, from 1.1.1.6, via Ethernet0/3

R1#show ip ospf topo path node-id 1.1.1.4 de
    Process OSPF-10

    Instance global
    Router ID : 1.1.1.1
    Area

```

```

Node 1.1.1.4  ABR, ASBR
  Algo 128, Distance 30
    Flags : Reachable

    1.2.0.2, from 1.1.1.2, via Ethernet0/1
      Weight : 0

  Algo 129, Distance 30
    Flags : Reachable

    1.1.1.6, from 1.1.1.6, via Ethernet0/0
      Weight : 0

    1.6.1.6, from 1.1.1.6, via Ethernet0/3
      Weight : 0

```

Use the command **show ip ospf topology route** display the path information of routes computed based on route calculation.

```
show ip ospf [process-id] [area-id] topology route [algorithmalgorithm-number]
[prefix/mask] [detail]
```

```

R1#show ip ospf topology route
Route Table of OSPF-10 with router ID 1.1.1.1 (VRF global)

1.1.1.4/32
  Algo 128, Metric 31, SID 132, Label 16132
    1.2.0.2, from 1.1.1.2, via Ethernet0/1
  Algo 129, Metric 31, SID 133, Label 16133
    1.1.1.6, from 1.1.1.6, via Ethernet0/0
    1.6.1.6, from 1.1.1.6, via Ethernet0/3
1.1.1.34/32
  Algo 128, Metric 21, SID 43, Label 16043
    1.2.0.2, from 1.1.1.2, via Ethernet0/1
1.1.1.45/32
  Algo 129, Metric 31, SID 4294967295, Label 1048577
    1.1.1.6, from 1.1.1.6, via Ethernet0/0
    1.6.1.6, from 1.1.1.6, via Ethernet0/3
1.1.1.54/32
  Algo 129, Metric 21, SID 45, Label 16045
    1.1.1.6, from 1.1.1.6, via Ethernet0/0
    1.6.1.6, from 1.1.1.6, via Ethernet0/3

R1#show ip ospf topo route 1.1.1.4/32 de
Route Table of OSPF-10 with router ID 1.1.1.1 (VRF global)

Route entry for 1.1.1.4/32
Priority : Medium
Flags : New

  Algo 128, Metric 31, SID 132, Label 16132
    Route type : Intra
    Flags: Inuse
    Prefix Contrib Algo 128 SID 132
    Total Metric : 31
    Contrib Flags : Inuse, Reachable
    SID Flags : Exp Null, PHP off, Index, Global, Node, Valid

    Path: 1.2.0.2, from 1.1.1.2, via Ethernet0/1
      Out Label : 16132
      Weight : 0

  Algo 129, Metric 31, SID 133, Label 16133

```

```
Route type : Intra
Flags: Inuse
Prefix Contrib Algo 129 SID 133
Total Metric : 31
Contrib Flags : Inuse, Reachable
SID Flags : Exp Null, PHP off, Index, Global, Node, Valid

Path: 1.1.1.6, from 1.1.1.6, via Ethernet0/0
Out Label   : 16133
Weight      : 0

Path: 1.6.1.6, from 1.1.1.6, via Ethernet0/3
Out Label   : 16133
Weight      : 0
```





## CHAPTER 18

# COE-PCE Initiated SR Policy with OSPF and IS-IS SR-TE Autoroute Announce

As part of a tactical TE solution, the Path Computation Element (PCE) can provision a Segment Routing Traffic Engineering (SR-TE) policy to mitigate link congestion.

Autoroute announcement is a steering mechanism in which IGP's automatically use the policy for destination's downstream of the policy end point. Autoroute announcement is performed using Cisco Crossworks Optimization Engine (COE). COE provides real-time network optimization allowing operators to maximize network utilization effectively and increase service velocity.

You can configure COE-PCE initiated SR policy in the following ways:

- PCE Initiated SR Policy with OSPF SR-TE Autoroute Announce — It enables a steering mechanism in which IGP's automatically use the SR-TE policy for destination's downstream of the policy end point.
- PCE-Initiated SR Policy with IS-IS SR-TE Autoroute Announce — It enables System-to-Intermediate System (IS-IS) interaction with traffic engineering to receive the SR-TE policies via autoroute announcement notifications.
- [COE-PCE Initiated SR Policy with OSPF Autoroute Announce, on page 219](#)
- [SR-PCE: Support for PCE-Initiated SR Policy and ISIS Autoroute, on page 226](#)

## COE-PCE Initiated SR Policy with OSPF Autoroute Announce

Table 14: Feature History

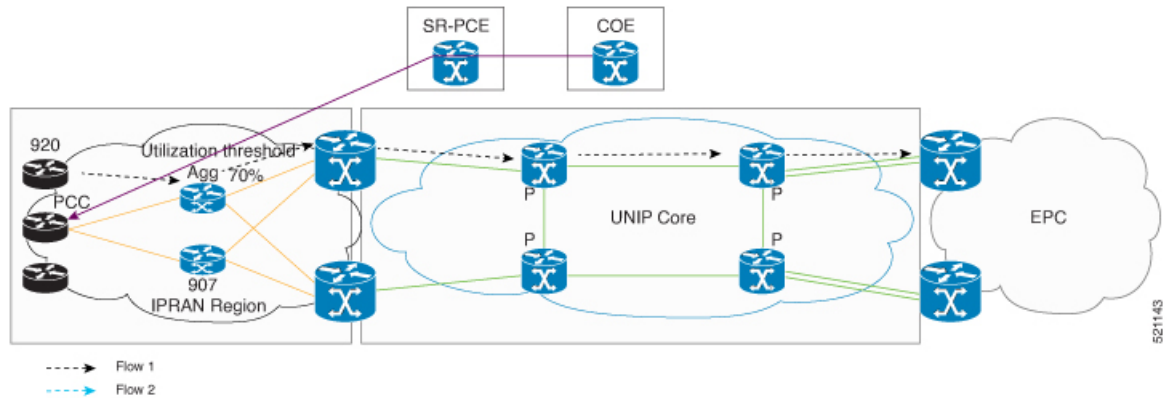
Feature Name	Release Information	Feature Description
PCE Initiated SR Policy with OSPF Autoroute Announce	Cisco IOS XE Bengaluru 17.4.1	This feature enables a steering mechanism in which IGP's automatically use the policy for destination's downstream of the policy end point.

A PCE collects various pieces of network information to determine traffic flows causing link congestion. The PCE computes a suitable path to divert those flows and to alleviate the congestion. The PCE then deploys the SR-TE policy to divert the traffic leading to the congestion using the Stateful Path Computation Element Protocol (PCEP) to provision the policy. When the congestion is alleviated, the SR-TE policy is removed.

The PCEP message contains SID list to be deployed by the head-end. Path Computation Client (PCC) profiles allow activation of autoroute announce for the policy provisioned by PCEP, using the profile IDs. The profile ID on the PCE and PCC should match, otherwise the policy is not provisioned. For example, if the PCE provisions a policy with profile ID 1 and the head-end where the policy is being provisioned also has the PCC profile ID 1 configured with autoroute announce, COE-PCE initiated SR policy is activated for that policy.

## COE-PCE Initiated SR Policy

Figure 25: COE-PCE Initiated SR Policy



The preceding topology shows how an SR-PCE policy is initiated from COE:

- SR policy is configured on the COE with profile ID.
- COE pushes the SR policy to PCE and PCE forwards the SR policy to PCC.
- Profile ID on PCC is matched with the profile ID on COE-PCE.
- OSPF autoroute announce is configured on the PCC.
- The policy gets provisioned.
- The data traffic now adheres to the SR policy that is pushed from the COE.
- Complete SR Policy manipulation occurs only on COE.

## Restrictions for PCE Initiated SR Policy

- A maximum of 500 SR policies are supported.
- Only native COE is supported.
- Effective Cisco IOS XE Bengaluru 17.5.1, Bandwidth optimization based on SR tactical policy is supported on RSP3.
- Bandwidth optimization by using COE is not supported.
- PIC core is not supported over SR-TE tunnel.
- PIC edge over SR-TE is not supported.
- Effective Cisco IOS XE Bengaluru 17.5.1, ECMP over SR-TE is supported on RSP3.

- 6PE and 6VPE are not supported with three and four transport labels.
- IPv6 is not supported.
- A maximum of 10,000 VPNv4 prefix limits are supported.
- BGP LU (RFC 3107) is not supported for intra-AS and inter-AS.

## ECMP Over SR-TE

*Table 15: Feature History*

Feature Name	Release Information	Feature Description
ECMP over SR-TE Policy	Cisco IOS XE Bengaluru 17.5.1	<p>This feature allows you to configure ECMP over SR-TE policies. In case of multiple paths, this feature enables mitigation of local congestion through load balancing.</p> <p>This feature is supported only on Cisco ASR 900 RSP3 module.</p>

The following sections explain how local congestion can be mitigated and how ECMP can be deployed over SR-TE policies to attain load balancing.



### Note

The traffic that is load balanced over multiple paths is HW-load balanced.

## Restrictions for ECMP over SR-TE Policies

Cisco ASR 900 RSP3 module supports **sr\_5\_label\_push\_enable** and **sr\_pfp\_enable** templates. Following restrictions apply for different template combinations.

With **sr\_5\_label\_push\_enable** template:

- Only one service label is supported with LB over SR-TE tunnels with three or four TE labels. This service label includes L3VPN, L2VPN, 6PE, 6VPE, and RFC 3107 BGP-LU label.
- 6PE and 6VPE are not supported with three and four SR-TE tunnel labels.
- Segment routing is not supported in **enable\_portchannel\_qos\_multiple\_active** template.
- HW load balancing for L2VPN/EVPN services is not supported if the L2VPN/EVPN destination has a static route configured over SR-TE tunnel.

With **sr\_pfp\_enable** template:

- SR PM HW time stamping is not supported.
- VLAN COS marking is not supported.
- HW load balancing is not supported.

- Policer based hierarchical QOS on the ingress is not supported.
- Short-Pipe tunneling mode is not supported.

Other Restrictions:

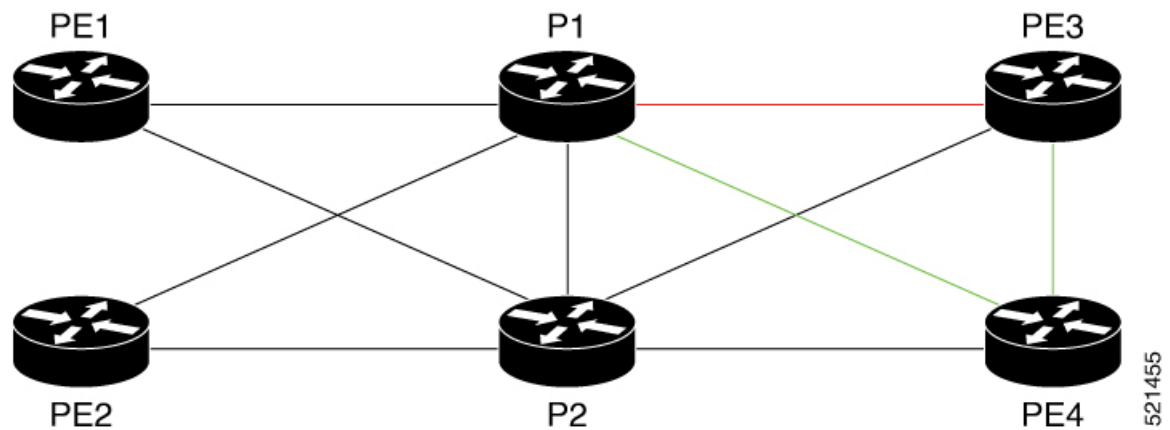
- PIC core over SR-TE tunnels are not supported.
- PIC edge over SR TE tunnels are not supported.
- PIC edge multipath over SR TE tunnels are not supported.
- W-ECMP is not supported.
- Next hop ECMP is not supported within an SR policy.
- Local congestion mitigation (LCM) is applicable only for best effort traffic. All other delay sensitive traffic uses safe SIDs (Flex Algo 128). Delay sensitive traffic is not redirected using the LCM tunnels.

## Local Congestion Mitigation

In today's network deployments it is important for every router in the network to have the capability to provision the traffic in such a way that it avoids the congestion based on the amount of traffic ingressing and egressing out of it. In order to provision this congestion mitigation, it is essential for the routers to support Equal Cost Multi-Path (ECMP) load balancing, that is, distributing the traffic based on the number of paths available to reach the destination.

Congestion mitigation helps the routers to move certain traffic to a different path than the current path, using the tactical SR policies. When the link congestion threshold is crossed, the COE (Cisco Optimization Engine) that monitors the link congestion based on the interface counters, pushes these tactical policies using PCE. These PCE initiated tactical policies that are used for local congestion mitigation (LCM) are deployed when necessary and only best effort traffic is load balanced over these tactical SR-TE policies.

**Figure 26: Illustration of Local Congestion Mitigation**



In the above topology, let us assume that the best effort traffic is coming in to P1 from PE1 and PE2 for the destination PE3 and the link between P1 and PE3 is congested. To mitigate the congestion between P1 and PE3, ECMP paths from P1 and PE3 are required. With segment routing this is achieved by deploying multiple tactical SR policies from P1 to PE3, one through directly connected link P1-PE3 and the other through the path P1-PE4-PE3. These policies are called tactical policies and are used to avoid local congestion mitigation by load balancing the best effort traffic over these tactical policies. The LCM is applicable only for best effort

traffic. All other delay sensitive traffic would use safe SIDs (Flex Algo 128). Delay sensitive traffic is not redirected using the LCM tunnels. Originating traffic is directed on non-LCM tunnels and transiting traffic with safe-SIDs is treated as normal label entry traffic and forwarded accordingly.

In the above topology, any node may deploy LCM tactical tunnels to mitigate congestion over a particular link. These nodes transit or sometimes originate the traffic to the LCM tunnel end points or even beyond the tunnel end points.

Let us assume that PE nodes originate the traffic and P nodes are transit node for the traffic originated somewhere else. Based on these combinations following are the different types of traffic that have to be considered:

As a PE Node,

- L3VPN best effort traffic
- L2VPN best effort traffic
- Global traffic

As a P node,

- Any traffic that comes in for a non-flexible algorithm 0 label is treated as an entry swap on the Label lookup.
- Any traffic that comes in for flexible algorithm 0 label is treated as a swap case or it may be translated to pop and push stack of labels, if there is an LCM created for that outgoing link based on congestion.

Based on the number of TE labels that the LCM tunnels have to push, the number of labels outside of TE labels can be either one or two (service labels).

## Load Balancing

At the head end, following are the different types of traffic that is subjected to load balancing. The traffic type here includes both best effort and delay sensitive.

As a PE Node,

- L3VPN traffic
- L2VPN traffic
- Global traffic

As a P node,

- Any traffic that comes in is treated based on the Label lookup.

## Autoroute Announcement

Autoroute announcement or bandwidth optimization is used to steer traffic away from congested links and better utilize the network.

The PCEP message contains SID list to be deployed by the head-end. Path Computation Client (PCC) profiles allow autoroute announce to be activated for the policy instantiated by PCEP, using the profile IDs. For example, if the PCE instantiates a policy with profile ID 1 and the head-end where the policy is being instantiated has the PCC profile ID 1 configured with autoroute announce, PCE initiated SR policy is activated for that policy.

Autoroute announce can be configured under both policies created with strict SID and policies created with non-strict SID. The main difference between configuring autoroute under policies created with strict SID (assume A) and non-strict SID (assume B) is that with A, the lookup entry will be programmed only in RIB whereas with B, the lookup entry will be programmed in RIB and LFIB for flexible algorithm label 0.

## Static Route Configuration

By adding a static route to the same destination but with different tunnels having the same endpoint, a load balancing is formed for the route over the tunnels configured. This is applicable for all types of traffic.

## Next Hop ECMP within a SR Policy

If there is a SR policy created to a destination with a set of SIDs and the SR policy headend have multiple equal paths to reach the next hop, no ECMP is formed to reach the next hop within the SR policy.

## Configuring ECMP over SR-TE Policy with OSPF Autoroute Announce

The below configuration shows how to configure ECMP over SR-TE policy with OSPF autoroute announce and PCE initiated segment routing policy with profile ID as 100.

```
pce
  address ipv4 13.13.13.13
  segment-routing traffic-eng
  peer ipv4 1.1.1.1
  segment-list name ssl
  index 1 mpls label 18021
  index 2 mpls label 18023
  policy 100
  binding-sid mpls 15999
  color 100 end-point ipv4 12.12.12.12
  candidate-paths
  preference 10
  explicit segment-list ssl
  !
  constraints
  segments
  dataplane mpls
  profile-id 100
```

Now, to push the PCE initiated OSPF autoroute announce from PCE to PCC, the profile IDs on PCE and PCC must match. The below configuration shows the PCC configuration and that the profile ID is matching with PCE and thus the autoroute announce is enabled. Based on the number of autoroute policies configured on the ECMP link, packets are load balanced on the ECMP links for the non-strict SIDs.

```
segment-routing traffic-eng
  pcc
  pce address 13.13.13.13 source-address 1.1.1.1
  profile 100
  autoroute
  include all
```

## Verifying SR Policy with Autoroute Announce

```
ASR903-R1#show segment-routing traffic-eng policy all

Name: *12.12.12.12|100 (Color: 100 End-point: 12.12.12.12)
Owners : PCEP
Status:
Admin: up, Operational: up for 66:41:16 (since 09-18 16:56:50.444)
Candidate-paths:
Preference 10 (PCEP):
```

```

PCC profile: 100
Dynamic (pce 13.13.13.13) (active)
Metric Type: TE, Path Accumulated Metric: 5
16003 [Prefix-SID, 3.3.3.3]
16012 [Prefix-SID, 12.12.12.12]
Attributes:
  Binding SID: 15999
Allocation mode: explicit
State: Programmed
Autoroute:
  Include all

```

## Verifying OSPF Autoroute for IGP

Use the following two commands to verify the OSPF Autoroute for IGP:

```

ASR903-R1#show ip cef 12.12.12.12 ----- IGP ROUTE
12.12.12.12/32
nexthop 12.12.12.12 Tunnel65536 ----- Tunnel pushed for IGP ROUTE

ASR903-R1# show ip cef 12.12.12.12 internal
12.12.12.12/32, epoch 3, RIB[I], refcnt 6, per-destination sharing
  sources: RIB
  feature space:
    IPRM: 0x00028000
    Broker: linked, distributed at 1st priority
    LFD: 12.12.12.12/32 0 local labels
        contains path extension list
  ifnums:
    Tunnel65536(64)
  path list 3C97B678, 3 locks, per-destination, flags 0x49 [shble, rif, hwc]
  path 3E393010, share 1/1, type attached nexthop, for IPv4
    MPLS short path extensions: [rib | lblmrg | srlbl] MOI flags = 0x1 label implicit-null

    nexthop 12.12.12.12 Tunnel65536, IP midchain out of Tunnel65536 2FFE3D00
  output chain:
    IP midchain out of Tunnel65536 2FFE3D00
    label [16012|16012]
    FRR Primary (0x3D9D4CE0)
      <primary: TAG adj out of Port-channel1, addr 100.0.0.2 3C9559C0>
      <repair: TAG adj out of BDI1110, addr 111.0.0.2 3C954FC0>

```

## Verify the Tunnel ID on the SR Policy

```

ASR903-R1# show segment-routing traffic-eng policy name margin detail
Name: Margin (Color: 1000 End-point: 12.12.12.12)
  Owners : CLI
  Status:
    Admin: up, Operational: up for 00:50:52 (since 09-16 11:00:06.697)
  Candidate-paths:
    Preference 10 (CLI):
      Dynamic (pce 13.13.13.13) (active)
        Metric Type: TE, Path Accumulated Metric: 5
        16012 [Prefix-SID, 12.12.12.12]
  Attributes:
    Binding SID: 15900
    Allocation mode: explicit
    State: Programmed
  IPv6 caps enabled
  Tunnel ID: 65536 (Interface Handle: 0x15B)
  Per owner configs:
    CLI
    Binding SID: 15900
  Stats:
    Packets: 535473 Bytes: 805338440

```

```

Event history:
Timestamp      Client      Event type      Context: Value
-----
09-16 11:00:06.377 CLI      Policy created   Name: CLI
09-16 11:00:06.418 CLI      Set colour       Colour: 1000
09-16 11:00:06.418 CLI      Set end point    End-point: 12.12.12.12
09-16 11:00:06.446 CLI      Set binding SID  BSID: Binding SID set
09-16 11:00:06.577 CLI      Set dynamic      Path option: dynamic
09-16 11:00:06.620 CLI      BSID allocated   FWD: label 15900
09-16 11:00:06.637 FH Resolution Policy state UP   Status: PATH RESOLVED
09-16 11:00:06.697 FH Resolution Policy state DOWN Status: PATH NOT RESOLVED
09-16 11:00:06.706 CLI      Set dynamic pce  Path option: dynamic pce
09-16 11:00:07.240 FH Resolution Policy state UP   Status: PATH RESOLVED
09-16 11:00:09.520 FH Resolution REOPT triggered Status: REOPTIMIZED

```

## SR-PCE: Support for PCE-Initiated SR Policy and ISIS Autoroute

Table 16: Feature History

Feature Name	Release Information	Description
SR-PCE: Support for PCE-Initiated SR Policy and ISIS Autoroute	Cisco IOS XE Cupertino 17.7.1	This feature enables System-to-Intermediate System (IS-IS) interaction with traffic engineering to receive the SR-TE policies via autoroute announcement notification. These notifications are used as IGP shortcuts during SPT computation and route calculation and are installed as nexthops for applicable routes in Routing Information Base (RIB) or MPLS Forwarding Infrastructure (MFI).

Prior to Cisco IOS XE Cupertino Release 17.7.1, PCE-initiated SR policy was only supported on OSPF protocol. Starting with Cisco IOS XE Cupertino Release 17.7.1, PCE-initiated SR policy is supported also on System-to-Intermediate System (IS-IS) that interacts with traffic engineering to receive the SR-TE policies via autoroute announcement notification. These notifications are used as IGP shortcuts during SPT computation and route calculation and are installed as nexthops for applicable routes in Routing Information Base (RIB) or MPLS Forwarding Infrastructure (MFI).

<need information to add network diagram>

## Configure PCE-Initiated SR Policy and ISIS Autoroute

To configure PCE-initiated SR policy and ISIS autoroute:

```

policy Margin
  color 1000 end-point 12.12.12.12
  binding-sid mpls 15900
  candidate-paths
  preference 10
  constraints

```



```

        segments
        dataplane mpls
    !
dynamic
    pcep
    metric
    margin
    absolute 5
segment-routing traffic-eng
                                pcc
pce address 13.13.13.13 source-address 1.1.1.1
profile 100
autoroute
include all

```

## Verification of IS-IS Autoroute Configuration

Use the **show isis segment-routing policy** command to verify the configuration of IS-IS segment routing policy.

Router2#**show isis segment-routing policy**

ISIS Router with ID (2.2.2.2) (Process ID 1)

Codes: SS - Strict SPF, SP - Default SPF  
r - relative, a - absolute, c - constant, n - none

Endpoint IP	Level/System-ID	Interface	Metric	Attributes	Last Updated
15.0.0.15	L1/R5.00	Tunnel65538	0	SS:n 00:01:14	00:01:14
16.0.0.16	L1/R6.00	Tunnel65539	-3	SS:r 00:00:03	00:01:21





## CHAPTER 19

# L2VPN over SR-TE Preferred Path

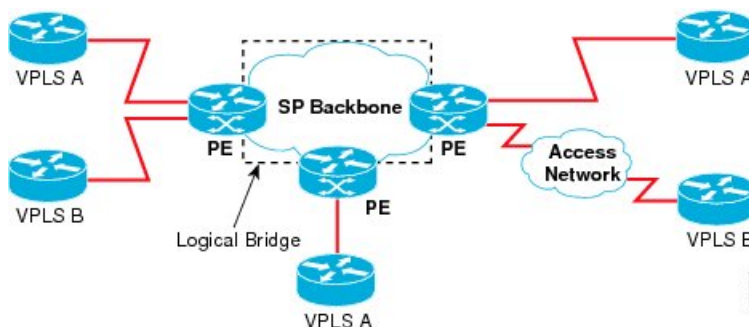
**Table 17: Feature History**

Feature Name	Release Information	Description
L2VPN over SR-TE Preferred Path	Cisco IOS XE Bengaluru 17.4.1	This feature allows you to configure an SR policy as the preferred path for a VPWS or VPLS pseudowire. VPWS or VPLS pseudowires between same PEs can be routed over different SR policies based on the requirements. Prior to this release, you could only steer the traffic using the SR policy for routing IPv4 traffic to a destination pseudowire (over IGP or BGP-LU).

Virtual Private LAN Services (VPLS) enables enterprises to link together multiple Ethernet-based LANs via the infrastructure provided by their service provider.

VPLS uses the service provider core to join multiple attachment circuits of an enterprise to simulate a virtual bridge. From the enterprise point of view, there is no topology for VPLS. All customer edge (CE) devices appear to connect to a logical bridge emulated by the service provider core.

**Figure 27: VPLS Topology**



Prior to Cisco IOS XE Bengaluru Release 17.4.1, L2VPN (VPLS or VPWS) traffic over SR policies could not be steered. You could only steer IPv4 traffic using the SR policy for routing IPv4 traffic to a destination pseudowire (over IGP or BGP-LU).

With Cisco IOS XE Bengaluru Release 17.4.1, you can now configure an SR policy as the preferred path for a VPWS or VPLS pseudowire. VPWS or VPLS pseudowires between same PEs can also be routed over different SR policies.



**Note** For SR labels till five, on the Cisco ASR 900 RSP3 module, you must use the **sr\_5\_label\_push\_enable** template to enable L2VPN over SR-TE preferred path.

### Disable Fallback Option

The disable fallback option disables the router from using the default path when the preferred path SR policy goes down.

- [Restrictions, on page 230](#)
- [Configuring L2VPN over SR-TE Preferred Path, on page 230](#)
- [Configuration Example: L2VPN over SR-TE Preferred Path, on page 231](#)
- [Verification of L2VPN over SR-TE Preferred Path Configuration, on page 232](#)

## Restrictions

- A traffic loss of approximately 10 seconds is observed while removing the preferred path configuration.
- NETCONF-YANG data model does *not* support template-based configuration.
- You cannot add On-Demand (ODN) policies to the preferred path.
- On the Cisco ASR 900 RSP3 module, you cannot use additional templates to support VPLS stats.
- The maximum number of SR labels supported is five. This includes one or two SR service labels.
- L2VPN over SR-TE preferred path is *only* supported on SR Per Destination Policy (PDP); and not on the SR Per-Flow Policy (PFP).
- L2VPN over SR-TE preferred path can *only* be configured using the pseudowire interface.

## Configuring L2VPN over SR-TE Preferred Path

You must create the SR static policy to configure L2VPN over SR-TE preferred path.

To create SR static policy for MPLS label:

```
configure terminal
segment-routing traffic-eng
segment-list name segment-name
index 1 mpls label first hop label
index 2 mpls label second hop label
!
policy policy-name
color color-code end-point destination IP Address
candidate-paths
preference preference
explicit segment-list segment-name
constraints
```

```
segments
dataplane mpls
```

You can also create SR static policy for the following:

- MPLS adjacency
- MPLS prefix

L2VPN over SR-TE preferred path can be configured in the following ways:

- Non-Template based Configuration
- Template-based Configuration

#### Non-template Based Configuration:

- **Create Pseudowire**

To create pseudowire:

```
interface pseudowire 1
encapsulation mpls
neighbor peer-address vc-id
```

- **Attach Policy Using Preferred Path**

To attach a policy using the preferred path:

```
interface pseudowire1
preferred-path segment-routing traffic-eng policy policy-name [disable-fallback]
```

#### Template-based Configuration:

- **Create Template Type Pseudowire**

To create template type pseudowire:

```
template type pseudowire name
encapsulation mpls
preferred-path segment-routing traffic-eng policy name [disable-fallback]
```

- **Attach Policy Using Preferred Path**

To attach a policy using the preferred path:

```
interface pseudowire 1
source template type pseudowire name
```

## Configuration Example: L2VPN over SR-TE Preferred Path

This example shows how to configure L2VPN over SR-TE preferred path.

```
!
segment-routing traffic-eng
segment-list name CE11-PE12-Seg
index 1 mpls label 16005
index 2 mpls label 16008
index 3 mpls label 16010
!
policy CE11-PE12
color 50 end-point 12.12.12.12
```

```

candidate-paths
  preference 100
  explicit segment-list Inter_IGP
  !
constraints
  segments
  dataplane mpls
!
interface pseudowire1000
  encapsulation mpls
  signaling protocol ldp
  neighbor 12.12.12.12 1000
  preferred-path segment-routing traffic-eng policy CE11-PE12
!
l2vpn vfi context VC_1000
  vpn id 1000
  member pseudowire1000
bridge-domain 1000
member GigabitEthernet0/1/0 service-instance 1000
member vfi VC_1000
!

```

## Verification of L2VPN over SR-TE Preferred Path Configuration

Use the **show segment-routing traffic-eng policy name *policy name* detail** command to verify the policy configuration:

```
Router#show segment-routing traffic-eng policy name CE11-PE12 detail
```

```

Name: CE11-PE12 (Color: 50 End-point: 12.12.12.12)
  Owners : CLI
  Status:
    Admin: up, Operational: up for 70:04:00 (since 08-17 07:55:36.536)
  Candidate-paths:
    Preference 100 (CLI):
      Explicit: segment-list IntraDomain (active)
      Weight: 1, Metric Type: TE
16005
16008
16010
Attributes:
  Binding SID: 20
  Allocation mode: dynamic
  State: Programmed
Tunnel ID: 65538 (Interface Handle: 0x20)
Per owner configs:
  CLI
  Binding SID: dynamic
Stats:
  Packets: 0 Bytes: 0

```

Event history:	Client	Event type	Context:
Timestamp			
Value	-----	-----	
-----: -----			
10-28 04:05:37.028	L2VPN	Policy created	Name: L2VPN
10-28 04:05:37.048	L2VPN	BSID allocated	FWD: label
20 10-28 04:05:37.494	L2VPN	Client removed	Owner:
Destroyed			

10-28 04:05:37.494	CLI	Set colour	Colour:
230			
10-28 04:05:37.494	CLI	Set end point	End-point:
12.12.12.12			
10-28 04:05:37.496	CLI	Set explicit path	Path option:
IntraDomain			
10-28 04:08:22.873	FH Resolution	Policy state UP	Status:
PATH RESOLVED			
10-28 04:08:45.630	FH Resolution	REOPT triggered	Status:
REOPTIMIZED			

Use **show mpls l2transport vc 1000 detail** command to verify the L2VPN over SR-TE preferred path:

```
Router#show mpls l2transport vc 1000 detail
Local interface: VFI VC_1000 vfi up
Interworking type is Ethernet
Destination address: 12.12.12.12, VC ID: 1000, VC status: up
Output interface: tu65538, imposed label stack {16005 16008 16010 32}
Preferred path: not configured
Default path: active
Next hop: 182.168.1.1
Create time: 1w4d, last status change time: 22:50:57
Last label FSM state change time: 22:51:46
Signaling protocol: LDP, peer 1.1.1.1:0 up
Targeted Hello: 2.2.2.2(LDP Id) -> 1.1.1.1, LDP is UP
Graceful restart: not configured and not enabled
Non stop routing: not configured and not enabled
Status TLV support (local/remote) : enabled/supported
LDP route watch : enabled
Label/status state machine : established, LruRru
Last local dataplane status rcvd: No fault
Last BFD dataplane status rcvd: Not sent
Last BFD peer monitor status rcvd: No fault
Last local AC circuit status rcvd: No fault
Last local AC circuit status sent: No fault
Last local PW i/f circ status rcvd: No fault
Last local LDP TLV status sent: No fault
Last remote LDP TLV status rcvd: No fault
Last remote LDP ADJ status rcvd: No fault
MPLS VC labels: local 26, remote 21
Group ID: local n/a, remote 16
MTU: local 9000, remote 9000
Remote interface description:
MAC Withdraw: sent:0, received:301
Sequencing: receive disabled, send disabled
Control Word: On (configured: autosense)
```







## CHAPTER 20

# SR-TE PM: Liveness of SR Policy Endpoint

**Table 18: Feature History**

Feature Name	Release Information	Description
SR-TE PM: Liveness of SR Policy Endpoint	Cisco IOS XE Bengaluru 17.5.1	<p>This feature enables Performance Measurement (PM) liveness detection and delay measurement for an SR policy on all the segment lists of every candidate path that are present in the forwarding table using PM probes. Thus, you can easily monitor the traffic path and efficiently detect any drop of traffic due to cable or hardware or configuration failures.</p> <p>This feature provides the following benefits:</p> <ul style="list-style-type: none"> <li>• End-to-end liveness is verified before activating the candidate path in the forwarding table.</li> <li>• End-to-end liveness failure can trigger re-optimization to another path by deactivating the current path.</li> </ul>

Metrics such as packet loss, delay, delay variation (jitter) and bandwidth utilization help you evaluate the performance of your network. You can use these metrics as input for Traffic Engineering (TE) and direct the flow of traffic through the network to conform to Service Level Agreements (SLAs). Network operators can use the performance measurement (PM) feature to monitor the network metrics for links. An SR-TE policy uses one or more candidate paths. A candidate path is a single segment list (SID-list) or a set of weighted SID-lists (for weighted equal cost multi-path [WECMP]). A candidate path is either dynamic or explicit.

Prior to Cisco IOS XE Bengaluru Release 17.5.1, PM could only measure delay for active candidate paths.

Starting with Cisco IOS XE Bengaluru Release 17.5.1, the liveness of SR policy endpoint configuration enables Performance Measurement (PM) liveness detection and delay measurement for an SR policy on all the segment lists of every candidate path that are present in the forwarding table using PM probes. Also,

SR-TE retains the existing traffic paths till PM liveness monitors and validates the new paths. Thus, you can easily monitor the traffic path and efficiently detect any drop of traffic due to cable or hardware or configuration failures.

PM sessions are created to the endpoint address of SR policy. If a candidate path has multiple segment-lists, PM sessions are created for each segment list. The probe query messages are sent using the segment-list SIDs (or labels) of the SR policy and SR-TE resolves the outgoing interface. SR-TE uses the First Hop Resolution (FHR) Equal Cost Multi Path Protocol (ECMP) feature.

Also, starting with this release, the default measurement mode for delay profile is changed from one-way mode to two-way mode.

**Note**

By default, a packet is sent every 3 seconds per session. With the default liveness detection multiplier, which is 3, a session goes down in approximately 9 seconds. The maximum PPS rate considered is approximately 30 PPS per session.

- [Benefits, on page 236](#)
- [Restrictions, on page 236](#)
- [Probe Query Message, on page 237](#)
- [SR-TE Re-optimization \(Make-Before-Break\), on page 237](#)
- [Performance Measurement Named Profile, on page 237](#)
- [Configure SR-TE PM: Liveness of SR Policy Endpoint, on page 237](#)
- [Verification of SR-TE PM: Liveness of SR Policy Endpoint Configuration, on page 240](#)
- [SR-PM Delay Deduction \(Loopback Mode\), on page 245](#)

## Benefits

This feature provides the following benefits:

- End-to-end liveness is verified before activating the candidate path in the forwarding table.
- End-to-end liveness failure can trigger re-optimization to another path by deactivating the current path.

## Restrictions

- A maximum of 500 SR-TE PM and 1000 SR-TE PM liveness session SR policies are supported for Cisco ASR RSP2 module and Cisco ASR RSP3 modules, respectively.
- A maximum of 10 CP per SR-TE policy is recommended.
- You must maintain a maximum of 500 (for Cisco ASR RSP2 module) and a maximum of 1000 (for Cisco ASR RSP3 module) PM PPS (PPS = PM Sessions X Burst Interval). The burst interval is 3 seconds by default.
- Liveness PCE-initiated policy is not supported.

## Probe Query Message

The probe messages defined in RFC5357 are used for delay measurement for links and end-to-end SR paths including SR policies. The probe query messages for performance measurement of an end-to-end SR Policy is sent using its SR-MPLS header containing the MPLS segment list. The sender IPv4 or IPv6 address is used as the source address. The endpoint IPv4 or IPv6 address is used as the destination address.

PM probes can allow both liveness monitoring and delay measurement using the same set of probes. This reduces the operational complexity in the network.

Starting with Cisco IOS XE Bengaluru Release 17.5.1, the default measurement mode for delay profile is changed from one-way mode to two-way mode. In two-way measurement mode, when using a bidirectional path, the probe response message as defined in Figure 6 is sent back to the sender node on the congruent path of the data traffic on the same reverse direction Link or associated reverse SR Policy.

### Probe Message Throughput

Probe messages run in asynchronous pipeline mode, where the querier does not wait for a reply message before sending the next query message. This allows to send probe queries at a faster rate and hence the liveness failure can be detected faster.

When a PM session starts, the query messages are received for that particular PM session. The liveness module waits until the pipe is full and then starts counting the missed consecutive packets after the. Once message counting has started, any delayed message (for example, due to queuing) is dropped.

## SR-TE Re-optimization (Make-Before-Break)

Prior to Cisco IOS XE Bengaluru Release 17.5.1, during SR-TE reoptimization, after each Path Calculation (PCALC), the algorithm to calculate the best path from the head end LSR to the tail end LSR, SR-TE used to destroy the existing traffic paths before forwarding the new ones.

Starting with Cisco IOS XE Bengaluru Release 17.5.1, SR-TE retains the existing traffic paths till PM liveness monitors and validates the new paths. Thus, you can easily monitor the traffic path and efficiently detect any drop of traffic due to cable or hardware or configuration failures.

## Performance Measurement Named Profile

Performance Measurement (PM) named profile provides an ability to create a number of profiles for a given transport mode (for example, interface or SR policy). The implementation provides only a global profile for each type of transport mode. This prevents from having different configuration for a transport mode. Since PM named profile enables the ability to have unique configuration for each instance of a transport mode, when a particular policy is configured with a particular configuration, rest of the policies also have the same configuration.

## Configure SR-TE PM: Liveness of SR Policy Endpoint

To configure SR-TE PM liveness:

**On Router 1:**

```

policy FAST
  color 100 end-point 11.11.11.11
  performance-measurement
    delay-measurement
      profile FAST
      liveness-detection
        invalidation-action down
  candidate-paths
    preference 100
    explicit segment-list SRTE1
    !
    constraints
      segments
        dataplane mpls
    !
  preference 2
  explicit segment-list SRTE11.11.11.11
  !
  constraints
    segments
      dataplane mpls

```

**On Router 2:**

```
Performance-measurement
```

## Configuration Example: Default Delay Profile for Liveness

### Liveness Detection Multiplier

Liveness detection multiplier is referred to the number of consecutive missed probe packets before PM session is declared as down. The default value is 3. The PM requires at least one probe packet to declare a PM session as us.

The following example shows the configuration of default delay profile for liveness:

#### 1. Profile Type for SR Policies Configuration:

```
Router#performance-measurement
delay-profile sr-policy
```

#### 2. SR-MPLS Policy Configuration:

```
probe
computation-interval [3-3600sec]
burst-interval [30-15000msec]
sweep
destination ipv4 127.x.x.x [1-128]
```

#### 3. Liveness Detection Multiplier Configuration:

```
liveness-detection multiplier [2-10]
```

## Configuration Example: PM Named Profiles

The following example shows the configuration of PM names profiles:

#### 1. SR Policy Delay Profile Configuration:

```
Router#performance-measurement
delay-profile sr-policy [policy name]
```

## 2. SR Policy Delay Profile Configuration Under Policy Bar:

```
Router#  
!  
segment-routing traffic-eng  
policy bar  
  performance-measurement  
  delay-measurement  
  profile [profile name]
```

## 3. Interface Delay Profile Configuration:

```
Router#performance-measurement  
  delay-profile interface [interface name]
```

## 4. Interface Delay Profile Configuration Under Interface Ethernet 0/0:

```
Router#performance-measurement  
  interface Ethernet 0/0  
  delay-measurement  
  profile [profile name]
```

**Note**

The name of the profile must be unique for a given transport mode and the same name can be used for different transport modes. Also, in absence of a profile, the probe will stop transmitting.

# Configuration Example: SR Policy for Liveness Detection

The following example shows the configuration of SR policy for liveness detection:

## 1. End-to-End Delay for All Candidate Paths:

```
Router# segment-routing traffic-eng  
policy [policy-name]  
color 4 endpoint ipv4 1.1.1.5  
performance-measurement  
delay-measurement
```

## 2. End-to-End Liveness Detection:

```
liveness-detection  
invalidation-action [Down|None]
```

Bidirectional Forwarding Detection (BFD) mechanisms are specified to monitor the unidirectional MPLS Label Switched Paths (LSPs) via BFD MPLS. Use the **invalidation-action** {down | none} command to set the action to be taken when BFD session is invalidated.

**Table 19: Syntax Description**

Keyword	Description
None	This is the default option. When the PM session goes down, use this option to record the failure without modifying the SR policy operational state.

Keyword	Description
Down	When the PM session liveness state is up, use this option to configure the LSP to operational up state. When the PM session goes down, the LSP state is immediately brought down.

## Verification of SR-TE PM: Liveness of SR Policy Endpoint Configuration

Use the **show segment-routing traffic-eng policy name [policy-name] detail** command to verify the SR-TE policy name configuration:

```
Router#show segment-routing traffic-eng policy name policy-name detail
Name: FAST (Color: 100 End-point: 11.11.11.11)
  Owners : CLI
  Status:
    Admin: up, Operational: up for 01:12:44 (since 01-11 17:17:00.092)
  Candidate-paths:
    Preference 100 (CLI):
      PM State: Up
      Explicit: segment-list SRTE1 (active)
        Weight: 1, Metric Type: TE
        16006 [Prefix-SID, 6.6.6.6]
        16008 [Prefix-SID, 8.8.8.8]
        16011 [Prefix-SID, 11.11.11.11]
    Preference 2 (CLI):
      PM State: Up
      Explicit: segment-list SRTE11.11.11.11 (inactive)
        Weight: 1, Metric Type: TE
        16011 [Prefix-SID, 11.11.11.11]
  Attributes:
    Binding SID: 16
    Allocation mode: dynamic
    State: Programmed
  Tunnel ID: 65536 (Interface Handle: 0x20)
  Per owner configs:
    CLI
    Binding SID: dynamic
  Stats:
    5 minute output rate 0 bits/sec, 0 packets/sec
    Packets: 0 Bytes: 0
  PM profile: FAST
Router#
Event history:
  Timestamp                Client      Event type      Context:
Value                      -----
-----
01-11 17:17:00.092        CLI        Policy created   Name: CLI
01-11 17:17:00.096        CLI        Set colour       Colour: 100
01-11 17:17:00.096        CLI        Set end point    End-point:
11.11.11.11
01-11 17:17:00.129        CLI        Set delay measure status:
Enabled
01-11 17:17:00.130        CLI        PM Profile name  config:
FAST
```

```

01-11 17:17:00.131      CLI      Set Live Detection      status:
Enabled
01-11 17:17:00.131      CLI      Set Live Invalidation   action:
down
01-11 17:17:00.134      CLI      Set explicit path       Path option:
SRTE1
01-11 17:17:00.139      CLI      Set explicit path       Path option:
SRTE11.11.11.11
01-11 17:17:11.033      CLI      BSID allocated          FWD: label
16
01-11 17:22:25.510      FH Resolution           Liveness                CP: 100,
SL1 is Waiting
01-11 17:24:27.038      PM              Liveness                CP: 100,
SL1 is Down
01-11 18:16:36.368      FH Resolution           Liveness                CP: 2, SL2
is Waiting
01-11 18:16:39.048      PM              Liveness                CP: 2, SL2
is Up
01-11 18:16:39.048      PM              Liveness                CP: 100,
SL1 is Up
01-11 18:16:39.053      FH Resolution           Policy state UP          Status:
PATH RESOLVED CP: 100

```

Use the **show performance-measurement summary** command to verify the performance measurement summary information:

```

Router#show performance-measurement summary
Total interfaces                : 0
Total SR Policies               : 1
Maximum PPS                    : 1000 pkts/sec

Interface Delay-Measurement:
Total sessions                  : 0
Counters:
  Packets:
    Total sent                  : 0
    Total received              : 0
  Errors:
    Total sent errors           : 0
    Total received errors       : 0
  Probes:
    Total started               : 0
    Total completed             : 0
    Total incomplete            : 0
    Total advertisements        : 0
SR Policy Delay-Measurement:
Total sessions                  : 2- 2 CP configured
Counters:
  Packets:
    Total sent                  : 2
    Total received              : 2
  Errors:
    Total sent errors           : 0
    Total received errors       : 0
  Probes:
    Total started               : 0
    Total completed             : 0
    Total incomplete            : 0
    Total advertisements        : 0
Global Delay Counters:
Total packets sent              : 2
Total query packets received    : 2
Total invalid session id        : 0
Total no session                : 0
HW Support for MPLS-GAL [RFC6374] timestamp : No

```

```

HW Support for IPv4 TWAMP [RF5357] timestamp : Yes
HW Support for IPv6 TWAMP [RF5357] timestamp : No
HW Support for 64 bit timestamp               : Yes
HW Support for IPv4 UDP Cheksum               : No

```

Use the **show performance-measurement sr-policy name** [*sr-policy name*] command to verify the performance measurement SR policy name configuration:

```

Router#show performance-measurement sr-policy name sr-policy name
SR Policy name: FAST
  Color                               : 100
  Endpoint                           : 11.11.11.11
  Source                             : 2.2.2.2
  Profile name                       : FAST
  Policy Update Timestamp             : 18:16:39  11 2021
  Number of candidate-paths          : 2

Candidate-Path:
  Preference                         : 2
  Protocol-origin                   : Configured
  Discriminator                     : 0
  Active:                           : No
  Number of segment-lists           : 1
  Number of atomic paths             : 1
  Number of live UP atomic paths: 1
  Number of live Unknown atomic : 0
Max Pkts per Burst                  : 1500
  Max Pkts per Probe                 : 15000
  AP Min Run per Probe               : 3
  Round-robin bursts                 : 1
  Round-robin probes                 : 1
Last advertisement:
  Advertised at: 18:18:56  11 2021 (1073 seconds ago)
  Advertised delays (uSec): avg: 520, min: 450, max: 893, variance: 70
Next advertisement:
  Check scheduled at the end of the current probe (roughly every 120 seconds)
  Aggregated delays (uSec): avg: 485, min: 421, max: 602, variance: 64
Last probe:
  Packets Sent: 10, received: 10
  Measured delays (uSec): avg: 488, min: 459, max: 550, variance: 29
Current probe:
  Packets Sent: 8, received: 8
  Measured delays (uSec): avg: 478, min: 447, max: 511, variance: 31

```

Use the **show performance-measurement sr-policy detail private verbose** command to verify the performance measurement SR policy detail configuration:

```

Router#show performance-measurement sr-policy detail private verbose
SR Policy name: FAST
  Color                               : 100----- policy color
  Endpoint                           : 11.11.11.11
  Source                             : 2.2.2.2
  Profile name                       : FAST ----- policy Name
  Policy Update Timestamp             : 22:31:38  12 2021
  Number of candidate-paths          : 2

Candidate-Path:
  Preference                         : 2
  Protocol-origin                   : Configured
  Discriminator                     : 0
  Active:                           : No
  Number of segment-lists           : 1
  Number of atomic paths             : 1
  Number of live UP atomic paths: 1
  Number of live Unknown atomic : 0
Max Pkts per Burst                  : 6

```



```

Max Pkts per Probe           : 60
AP Min Run per Probe        : 3
Round-robin bursts          : 1
Round-robin probes          : 1
Last advertisement:
  Advertised at: 11:41:37 13 2021 (1879 seconds ago)
  Advertised delays (uSec): avg: 500, min: 435, max: 924, variance: 65
Next advertisement:
  Check scheduled in 1 more probe (roughly every 120 seconds)
  Aggregated delays (uSec): avg: 702, min: 437, max: 3940, variance: 265
Last probe:
  Packets Sent: 10, received: 10
  Measured delays (uSec): avg: 527, min: 441, max: 686, variance: 86
Current probe:
  Packets Sent: 7, received: 7
  Measured delays (uSec): avg: 452, min: 436, max: 475, variance: 16
Segment-List:
  Name                       : SL2
  Number of atomic paths     : 1
  Last advertisement:
    Advertised at: 11:41:37 13 2021 (1879 seconds ago)
    Advertised delays (uSec): avg: 500, min: 435, max: 924, variance: 65
  Next advertisement:
    Aggregated delays (uSec): avg: 702, min: 437, max: 3940, variance: 265
  Last probe:
    None
Current probe:
  None
Atomic path:
  Hops                       : 11.11.11.11
  Labels                     : 16011
  Outgoing Interface         : GigabitEthernet0/5/1
  Next Hop                   : 102.0.0.1
  Destination                 : 11.11.11.11
  Session ID                 : 409
  Last advertisement:
    Advertised at: 11:41:37 13 2021 (1879 seconds ago)
    Advertised reason: Periodic timer, avg delay threshold crossed
    Advertised delays (uSec): avg: 500, min: 435, max: 924, variance: 61
  Next advertisement:
    Aggregated delays (uSec): avg: 702, min: 437, max: 3940, variance: 263
    Rolling average (uSec): 625
  Last probe:
    None
Current probe:
  Packets Sent: 7, received: 7
  Measured delays (uSec): avg: 452, min: 436, max: 475, variance: 16
Probe samples:
  Packet Rx Timestamp Measured Delay
  12:12:55 13 2021 441680
  12:12:52 13 2021 475040
  12:12:49 13 2021 436400
  12:12:46 13 2021 438320
  12:12:43 13 2021 443120
  12:12:40 13 2021 471200
  12:12:37 13 2021 461800
Liveness Detection:
  Session Creation Timestamp: 22:31:34 12 2021
  Session State: Up
  Last State Change Timestamp: 22:31:35 12 2021
  Missed count [consecutive]: 0
  Received count [consecutive]: 16427
  Backoff : 0
  Unique Path Name : Path-3

```

```

Candidate-Path:
  Preference                : 100
  Protocol-origin           : Configured
  Discriminator             : 0
  Active:                   : Yes
  Number of segment-lists   : 1
  Number of atomic paths    : 1
  Number of live UP atomic paths: 1
  Number of live Unknown atomic : 0
  Max Pkts per Burst       : 6
  Max Pkts per Probe       : 60
  AP Min Run per Probe     : 3
  Round-robin bursts       : 1
  Round-robin probes       : 1
  Last advertisement:
    Advertised at: 01:35:36 13 2021 (38240 seconds ago)
    Advertised delays (uSec): avg: 507, min: 430, max: 1053, variance: 77
  Next advertisement:
    Check scheduled in 1 more probe (roughly every 120 seconds)
    Aggregated delays (uSec): avg: 533, min: 443, max: 846, variance: 90
  Last probe:
    Packets Sent: 10, received: 10
    Measured delays (uSec): avg: 541, min: 443, max: 846, variance: 98
Current probe:
  Packets Sent: 7, received: 7
  Measured delays (uSec): avg: 478, min: 444, max: 502, variance: 34

Segment-List:
  Name                      : SL1
  Number of atomic paths    : 1
  Last advertisement:
    Advertised at: 01:35:36 13 2021 (38240 seconds ago)
    Advertised delays (uSec): avg: 507, min: 430, max: 1053, variance: 77
  Next advertisement:
    Aggregated delays (uSec): avg: 533, min: 443, max: 846, variance: 90
  Last probe:
    None
  Current probe:
    None
Atomic path:
  Hops                      : 6.6.6.6, 8.8.8.8, 11.11.11.11
  Labels                    : 16006, 16008, 16011
  Outgoing Interface       : GigabitEthernet0/5/1
  Next Hop                  : 102.0.0.1
  Destination               : 11.11.11.11
  Session ID                : 408
Candidate-Path:
  Preference                : 100
  Protocol-origin           : Configured
  Discriminator             : 0
  Active:                   : Yes
  Number of segment-lists   : 1
  Number of atomic paths    : 1
  Number of live UP atomic paths: 1
  Number of live Unknown atomic : 0
  Max Pkts per Burst       : 6
  Max Pkts per Probe       : 60
  AP Min Run per Probe     : 3
  Round-robin bursts       : 1
  Round-robin probes       : 1
  Last advertisement:
    Advertised at: 01:35:36 13 2021 (38240 seconds ago)
    Advertised delays (uSec): avg: 507, min: 430, max: 1053, variance: 77
  Next advertisement:

```

```

Check scheduled in 1 more probe (roughly every 120 seconds)
Aggregated delays (uSec): avg: 533, min: 443, max: 846, variance: 90
Last probe:
  Packets Sent: 10, received: 10
  Measured delays (uSec): avg: 541, min: 443, max: 846, variance: 98
Last advertisement:
  Advertised at: 01:35:36 13 2021 (38240 seconds ago)
  Advertised reason: Periodic timer, avg delay threshold crossed
  Advertised delays (uSec): avg: 507, min: 430, max: 1053, variance: 60
Next advertisement:
  Aggregated delays (uSec): avg: 533, min: 443, max: 846, variance: 85
  Rolling average (uSec): 533
Last probe:
  None
Current probe:
  Packets Sent: 7, received: 7
  Measured delays (uSec): avg: 478, min: 444, max: 502, variance: 34
Probe samples:
Packet Rx Timestamp Measured Delay
12:12:56 13 2021 491200
12:12:53 13 2021 488440
12:12:50 13 2021 444280
12:12:47 13 2021 466360
12:12:44 13 2021 455160
12:12:41 13 2021 502480
12:12:38 13 2021 500720
Liveness Detection:
  Session Creation Timestamp: 22:31:34 12 2021
  Session State: Up
  Last State Change Timestamp: 22:31:35 12 2021
  Missed count [consecutive]: 0
  Received count [consecutive]: 16427
  Backoff : 0
  Unique Path Name : Path-4

```

## SR-PM Delay Deduction (Loopback Mode)

**Table 20: Feature History**

Feature Name	Release Information	Description
SR-PM Delay Deduction (Loopback Mode)	Cisco IOS XE Bengaluru 17.5.1	This feature improves the SR-PM detection time as the PM probes are not punted on the remote nodes. Also, it does not require a third-party support for interoperability.

Prior to Cisco IOS XE Bengaluru Release 17.5.1, you could not configure loopback mode for SR-PM detection.

Starting with Cisco IOS XE Bengaluru Release 17.5.1, you can configure the loopback mode at the head nodes of the router while the remote nodes can be used for any third-party configuration. As the PM probes are not punted on the remote nodes, the SR-PM detection time is enhanced and improved. Also, third-party support for interoperability is not required.

## Configuration Example: SR-PM Delay Deduction (Loopback Mode)

Use the **reverse-path** command to configure reverse path configuration before you configure loopback mode. This configuration enables an MPLS label configuration that is pushed in PM probe messages above the existing label list received from the SR policy manager for the particular path.

The following example shows the configuration of SR-PM delay deduction (loopback mode):

```
policy FAST
  color 100 end-point 11.11.11.11
  performance-measurement
    delay-measurement
      profile FAST
      liveness-detection
      invalidation-action down
      reverse-path
        label 16002
  candidate-paths
    preference 100
    explicit segment-list SRTE1
    !
    constraints
      segments
        dataplane mpls
    !
    !
  preference 2
  explicit segment-list SRTE11.11.11.11
  !
  constraints
    segments
      dataplane mpls
  !
  !
  !

performance-measurement
!
delay-profile sr-policy name FAST
probe
  measurement-mode loopback
```

## Verify SR-PM Delay Deduction (Loopback Mode) Configuration

Use the **show performance-measurement profile sr-policy *policy-name*** command to verify the performance measurement SR policy configuration using the loopback mode:

```
Router#show performance-measurement profile sr-policy policy-name
FAST SR Policy Delay Measurement:
Profile configuration:
  Measurement Type                : Loopback
  Computation interval             : 30 (effective : 30) seconds
  Burst interval                   : 3000 mSec
  Burst count                      : 10
  Protocol                        : TWAMP-Lite Unauth
  Destination sweeping mode       : Disabled
  Periodic advertisement          : Enabled
    Interval                      : 120 (effective: 120) sec
    Threshold                     : 10%
    Minimum-Change                 : 500 uSec
  Accelerated advertisement       : Disabled
```

```

Threshold crossing check           : Average-delay
Liveness-detection multiplier     : 3

```

Use the **show segment-routing traffic-eng policy name** *[policy-name]* **detail** command to verify the SR-TE policy name configuration using loopback mode:

```

Router#show segment-routing traffic-eng policy name policy-name detail
Name: FAST (Color: 100 End-point: 11.11.11.11)
  Owners : CLI
  Status:
    Admin: up, Operational: up for 183:26:26 (since 01-28 20:10:24.628)
  Candidate-paths:
    Preference 100 (CLI):
      PM State: Up
      Explicit: segment-list SRTE1 (active)
        Weight: 1, Metric Type: TE
        16006 [Prefix-SID, 6.6.6.6]
        16008 [Prefix-SID, 8.8.8.8]
        16011 [Prefix-SID, 11.11.11.11]
    Preference 2 (CLI):
      PM State: Up
      Explicit: segment-list SRTE11.11.11.11 (inactive)
        Weight: 1, Metric Type: TE
        16011 [Prefix-SID, 11.11.11.11]
  Attributes:
    Binding SID: 16
    Allocation mode: dynamic
    State: Programmed
  Tunnel ID: 65536 (Interface Handle: 0x20)
  Per owner configs:
    CLI
    Binding SID: dynamic
  Stats:
    5 minute output rate 0 bits/sec, 0 packets/sec
    Packets: 0 Bytes: 0
  PM profile: FAST

```

Event history:	Timestamp	Client	Event type	Context:
Value	-----	-----	-----	-----:
-----				
01-20 08:50:22.483	FH Resolution	Liveness	CP: 2, SL2	
is Waiting				
01-20 08:50:22.492	PM	Liveness	CP: 100,	
SL1 is Up				
01-20 08:50:22.492	PM	Liveness	CP: 2, SL2	
is Up				
01-20 08:50:22.494	FH Resolution	REOPT triggered	Status:	
REOPTIMIZED CP: 100				
01-20 08:51:54.426	FH Resolution	REOPT triggered	Status:	
REOPTIMIZED CP: 100				
01-20 08:52:00.964	FH Resolution	REOPT triggered	Status:	
REOPTIMIZED CP: 100				
01-20 08:55:10.264	FH Resolution	REOPT triggered	Status:	
REOPTIMIZED CP: 100				
01-20 12:04:06.663	FH Resolution	Liveness	CP: 100,	
SL1 is Waiting				
01-20 12:04:06.664	FH Resolution	Liveness	CP: 2, SL2	
is Waiting				
01-20 12:04:08.773	PM	Liveness	CP: 2, SL2	
is Up				
01-20 12:04:08.836	FH Resolution	REOPT triggered	Status:	
REOPTIMIZED CP: 2				

## Verify SR-PM Delay Deduction (Loopback Mode) Configuration

01-20 12:04:08.954	PM	Liveness	CP: 100,
SL1 is Up			
01-20 12:04:09.001	FH Resolution	REOPT triggered	Status:
REOPTIMIZED CP: 100			
01-20 14:31:41.138	FH Resolution	Liveness	CP: 100,
SL1 is Waiting			
01-20 14:31:41.138	FH Resolution	Liveness	CP: 2, SL2
is Waiting			
01-20 14:31:44.292	PM	Liveness	CP: 2, SL2
is Up			
01-20 14:31:44.430	FH Resolution	REOPT triggered	Status:
REOPTIMIZED CP: 2			
01-20 14:31:45.099	PM	Liveness	CP: 100,
SL1 is Up			
01-20 14:31:45.232	FH Resolution	REOPT triggered	Status:
REOPTIMIZED CP: 100			
01-21 16:17:05.118	FH Resolution	REOPT triggered	Status:
REOPTIMIZED CP: 100			
01-21 16:58:34.741	FH Resolution	REOPT triggered	Status:
REOPTIMIZED CP: 100			
01-22 08:13:05.533	FH Resolution	REOPT triggered	Status:
REOPTIMIZED CP: 100			
01-22 08:17:51.393	FH Resolution	REOPT triggered	Status:
REOPTIMIZED CP: 100			
01-22 13:10:38.098	FH Resolution	REOPT triggered	Status:
REOPTIMIZED CP: 100			
01-28 12:04:30.402	FH Resolution	REOPT triggered	Status:
REOPTIMIZED CP: 100			
01-28 12:07:29.883	FH Resolution	Liveness	CP: 100,
SL1 is Waiting			
01-28 12:07:29.883	FH Resolution	Liveness	CP: 2, SL2
is Waiting			
01-28 12:07:29.961	PM	Liveness	CP: 100,
SL1 is Up			
01-28 12:07:29.962	PM	Liveness	CP: 2, SL2
is Up			
01-28 12:07:30.323	FH Resolution	REOPT triggered	Status:
REOPTIMIZED CP: 100			
01-28 12:44:13.208	FH Resolution	Policy state DOWN	Status:
PATH NOT RESOLVED			
01-28 12:44:13.392	PM	Liveness	CP: 100,
SL1 is Unknown			
01-28 12:44:13.392	PM	Liveness	CP: 2, SL2
is Unknown			
01-28 19:40:14.414	FH Resolution	Liveness	CP: 2, SL2
is Waiting			
01-28 19:40:16.137	PM	Liveness	CP: 2, SL2
is Up			
01-28 19:40:16.277	FH Resolution	Policy state UP	Status:
PATH RESOLVED CP: 2			
01-28 20:10:24.628	FH Resolution	Policy state DOWN	Status:
PATH NOT RESOLVED			
01-28 20:10:24.971	PM	Liveness	CP: 2, SL2
is Unknown			
01-28 20:10:27.656	FH Resolution	Liveness	CP: 2, SL2
is Waiting			
01-28 20:10:30.219	PM	Liveness	CP: 2, SL2
is Up			
01-28 20:10:30.311	PM	Liveness	CP: 2, SL2
is Unknown			
02-05 11:27:57.404	CLI	Set PM delay loopbac	:
02-05 11:34:48.918	FH Resolution	Liveness	CP: 100,
SL1 is Waiting			
02-05 11:35:42.025	FH Resolution	Liveness	CP: 2, SL2

```

is Waiting
02-05 11:35:44.792          PM          Liveness          CP: 100,
SL1 is Up
02-05 11:35:45.030          PM          Liveness          CP: 2, SL2
is Up
02-05 11:35:45.031          FH Resolution    Policy state UP    Status:
PATH RESOLVED CP: 2
02-05 11:35:46.701          FH Resolution    REOPT triggered    Status:
REOPTIMIZED CP: 2
02-05 11:35:47.937          PM          Liveness          CP: 100,
SL1 is Up
02-05 11:35:47.938          FH Resolution    REOPT triggered    Status:
REOPTIMIZED CP: 100

```

Use the **show performance-measurement counters sr-policy name** *sr-policy name* to display the PM link-delay session counters.

Router#**show performance-measurement counters sr-policy name** *sr-policy name*

SR Policy name: FAST

Candidate-Path:

```

Preference          : 2
Protocol-origin      : Configured
Discriminator        : 0
Active              : No

```

Packets:

```

Total sent          : 55
Total received      : 55

```

Errors:

```

Total sent errors    : 0
Total received errors : 0

```

Probes:

```

Total started        : 5
Total completed      : 5
Total incomplete     : 0
Total advertisements : 1

```

Segment-list:

```

Name                : SL2

```

Packets:

```

Total sent          : 55
Total received      : 55

```

Errors:

```

Total sent errors    : 0
Total received errors : 0

```

Probes:

```

Total started        : 5
Total completed      : 5
Total incomplete     : 0
Total advertisements : 1

```

Candidate-Path:

```

Preference          : 100
Protocol-origin      : Configured
Discriminator        : 0
Active              : Yes

```

Packets:

```

Total sent          : 56
Total received      : 56

```

Errors:

```

Total sent errors    : 0
Total received errors : 0

```

Probes:

```

Total started        : 5
Total completed      : 5
Total incomplete     : 0
Total advertisements : 0

```

```
Segment-list:
  Name                               : SL1
  Packets:
    Total sent                       : 56
    Total received                   : 56
  Errors:
    Total sent errors                : 0
    Total received errors            : 0
  Probes:
    Total started                    : 5
    Total completed                  : 5
    Total incomplete                 : 0
    Total advertisements              : 0
```





## CHAPTER 21

# EVPN Integrated Routing and Bridging (L2 and L3 Anycast Gateway) and Data Center Interconnect or Border Leaf (Single Homing)

Table 21: Feature History

Feature Name	Release Information	Description
EVPN Integrated Routing and Bridging (L2 and L3 Anycast Gateway) and Data Center Interconnect or Border Leaf (Single Homing)	Cisco IOS XE Bengaluru 17.5.1	<p>This feature allows the devices to forward both layer 2 or bridged and layer 3 or routed traffic providing optimum unicast and multicast forwarding for both intra-subnets and inter-subnets within and across data centers. Data Center Interconnects (DCI) products are targeted at the Edge or Border Leaf (BL) of data center environments, joining data centers to each other in a point-to-point or point-to-multipoint fashion, or at times extending the connectivity to internet gateways or peering points.</p> <p>This feature is supported on Cisco NCS 4201/4202 routers.</p>

Prior to Cisco IOS XE Bengaluru Release 17.5.1, EVPN over MPLS network could only allow the routers to forward layer 2 traffic.

Starting with Cisco IOS XE Bengaluru Release 17.5.1, this feature allows the router in an EVPN over MPLS network to perform both bridging and routing. Integrated Routing and Bridging (IRB) provides the ability to route between a bridge group and a routed interface using a BVI. The BVI is a virtual interface within the router that acts like a normal routed interface. A BVI is associated with a single bridge domain and represents the link between the bridging and the routing domains on the router. An EVPN-based IRB allows the routers to forward both Layer 2 or bridged and Layer 3 or routed traffic providing optimum unicast and multicast forwarding for both intra-subnets and inter-subnets within and across data centers. Data Center Interconnects (DCI) products are targeted at the Edge or Border leaf (BL) of data center environments, joining data centers

to each other in a point-to-point or point-to-multipoint fashion, or at times extending the connectivity to internet gateways or peering points.

A bridge domain performs bridging when it forwards traffic to the same subnet. Similarly, a bridge domain interface performs routing when it forwards traffic to a different subnet. The devices in the network forward traffic to each other through the Distributed Anycast Gateways (DAG). The Ethernet VPN over MPLS Integrated IRB Single-Homing (SH) with DAG feature provides support for symmetric IRB model.

In symmetric IRB, both the ingress and egress bridge domain interfaces perform both bridging and routing. A packet first moves through a MAC VRF followed by an IP VRF of the ingress device. It then moves through an IP VRF followed by a MAC VRF on the PE of the egress device. The PEs of ingress and egress devices equally share all the packet processing associated with intersubnet forwarding semantics.

In symmetric IRB, you are required to define only the endpoints on the ingress and egress bridge domain interfaces. Symmetric IRB offers better scalability with the BGP EVPN over MPLS fabric.

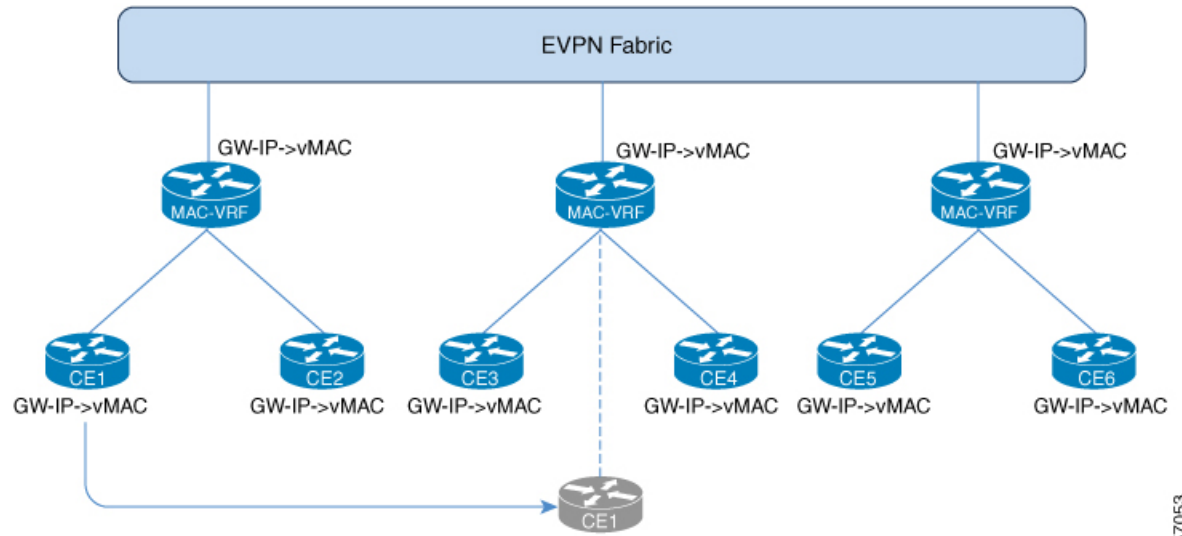
- [Distributed Anycast Gateway \(DAG\) with Bridge Domains, on page 252](#)
- [Symmetric IRB with MPLS on Distributed Gateways, on page 253](#)
- [Host MAC-IP Binding on a Single-Homed DAG, on page 254](#)
- [Host MAC-IP Mobility, on page 254](#)
- [ARP and ND Flooding Suppression, on page 255](#)
- [Prerequisites, on page 255](#)
- [Restrictions, on page 256](#)
- [Configuring EVPN Integrated Routing and Bridging \(L2 and L3 Anycast Gateway\) and Data Center Interconnect or Border Leaf \(Single Homing\), on page 256](#)
- [Verification of EVPN Integrated Routing and Bridging \(L2 and L3 Anycast Gateway\) and Data Center Interconnect or Border Leaf \(Single Homing\) Configuration, on page 258](#)
- [EVPN-IRB DHCP v4 and v6 Relay over Segment Routing, on page 264](#)
- [Stitching of Subnet Route from EVPN to L3VPN, on page 271](#)

## Distributed Anycast Gateway (DAG) with Bridge Domains

Distributed Anycast Gateway (DAG) is a default gateway addressing mechanism in a BGP EVPN fabric. The feature enables the use of the same gateway IP and MAC address across all the devices in an EVPN over MPLS network with IRB. This ensures that every device functions as the default gateway for the workloads directly connected to it. The feature facilitates flexible workload placement, host mobility, and optimal traffic forwarding across the BGP EVPN fabric.

In the scenario below, the DAGs are directly attached to hosts or network with IP-VRF routing enabled on the IRB (BDI) interfaces on the gateways. To reduce the complexity, only virtual MAC DAGs are supported and the Duplication Address Detection (DAD) for IPv6 on the BDI interfaces on DAG is disabled.

Figure 28: DAG with Bridge Domains



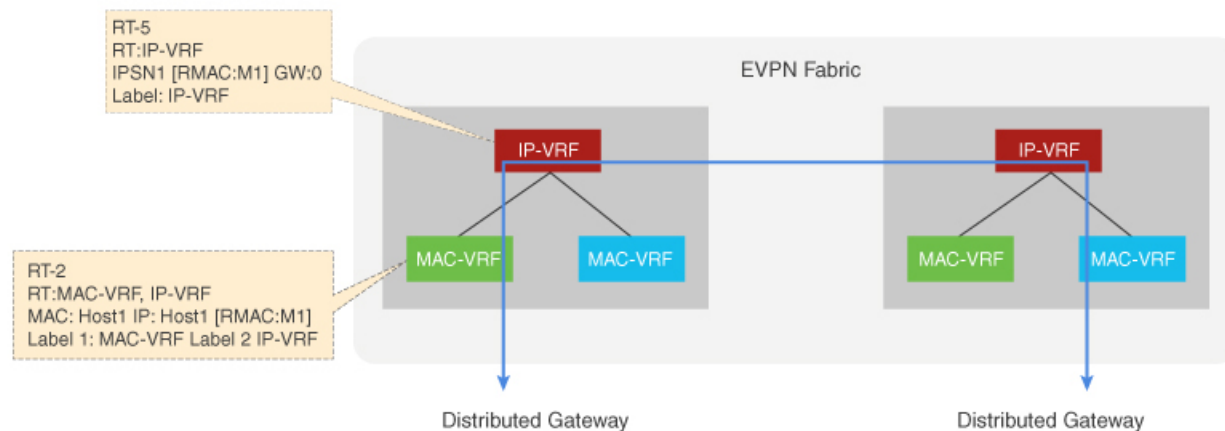
On the DAG, the bridge domain checks if an Address Resolution Protocol (ARP) or Neighbor Discovery Protocol (NDP) packet from a local host is sent to the BDI (Gateway) IP addresses. If the packet is sent to BDI (Gateway) IP addresses, this packet is handled by local BDI and it is not flooded into the bridge domain and sent across the EVPN IRB fabric.

## Symmetric IRB with MPLS on Distributed Gateways

Symmetric IRB is a distributed routing model which utilizes direct IP-VRF to IP-VRF connectivity for inter-subnet traffic. To support symmetric IRB, the native IRB needs to be enabled on distributed gateways by creating the BDIs, configuring virtual MAC, IP-VRF, and anycast IP address.

After the native IRB is enabled, BGP allocates the L3 label for the RT-2's and RT-5's per VRF basis and advertises it.

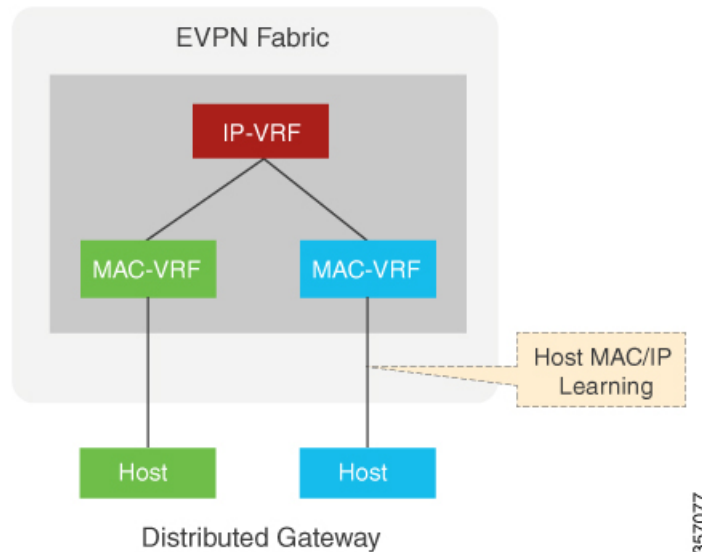
Figure 29: Symmetric IRB with MPLS on Distributed Gateways



## Host MAC-IP Binding on a Single-Homed DAG

The Host MAC-IP binding is learned by snooping Address Resolution Protocol (ARP), Neighbor Discovery Protocol, or DHCP packets. After the MAC-IP binding is learned, an age timer (AGE\_TIME) is applied to the locally learned binding entry. The binding entry is refreshed whenever the host initiates ARP or ND procedures.

**Figure 30: Host MAC-IP Binding on a Single-Homed DAG**



357077

## Host MAC-IP Mobility

The host MAC-IP mobility helps to handle the following events:

- Host Move Learn from Data Packet and Generic Attribute Registration Protocol (GARP)
- Host Move Detection for Silent Host

Also, the host MAC-IP mobility supports the following scenarios:

- Moving MAC from local to local
- Moving MAC from local to remote
- Moving MAC from remote to local
- Moving IP local to local
- Moving IP from local to remote
- Moving IP from remote to local

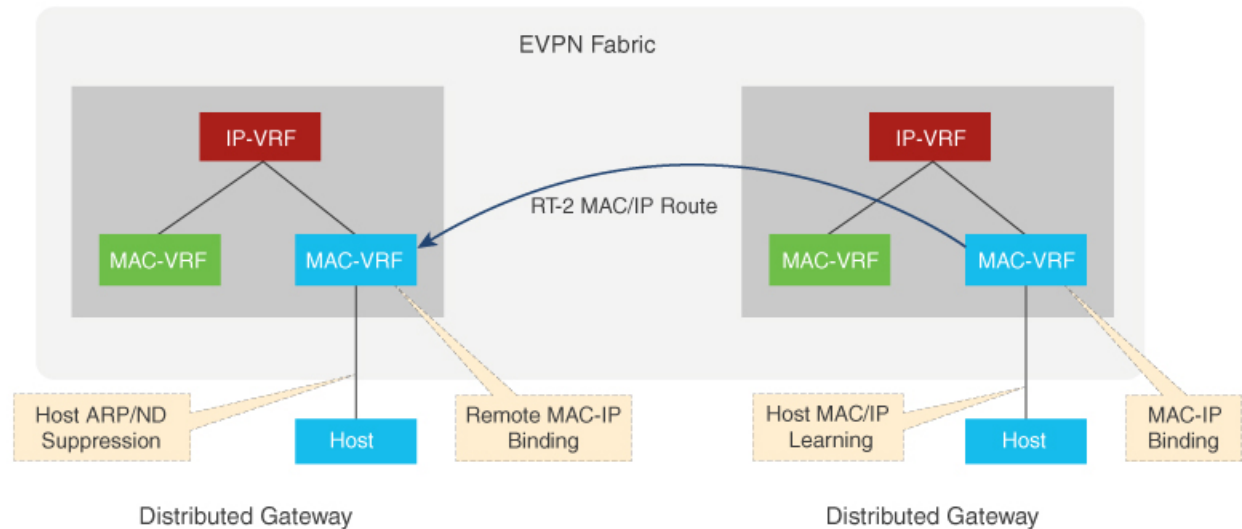
# ARP and ND Flooding Suppression

The Address Resolution Protocol (ARP) and Neighbor Discovery (ND) Protocol flooding suppression depend on device tracking enabled on the same VLAN or interface. The Switch Integrated Security Feature based (SISF-based) device tracking helps to track the presence, location, and movement of end nodes in the network. The device receives the SISF snoops traffic that extracts the device identity (MAC and IP address) and stores it in a binding table. SISF-based device tracking supports both IPv4 and IPv6.

When you enable IPv4 or IPv6 flooding suppression, it helps to minimize the flooding of a broadcast or multicast packet over the EVPN IRB fabric and to remote CEs such as host and router. The multicast and broadcast suppression capabilities help to preserve bandwidth in wireless networks.

This helps to suppress the broadcast (ARP) or link-local multicast (NDP) messages circulating in the layer 2 domain and the packets are relayed after converting their L2 addresses to unicast.

**Figure 31: ARP and ND Flooding Suppression**



## Prerequisites

- Host MAC-IP learning
- Symmetric IRB for IP-VRF to IP-VRF inter-subnet traffic over MPLS
- DAG with bridge-domain
- Host MAC-IP mobility
- ARP/ND flooding suppression

## Restrictions

- DHCP Snooping and EVPN IRB snooping cannot be enabled on the same bridge domain.
- Stitching on a collapsed Border Leaf (BL) or Spine Leaf is *not* supported.
- EVPN IRB feature and Hot Standby Router Protocol (HSRP) or Virtual Router Redundancy Protocol (VRRP) feature are *not* supported simultaneously on a router.
- EVPN Layer 2 is *not* supported on Media Access Control Security (MACsec).
- Asymmetric IRB and centralized IRB are *not* supported.
- Only Virtual MAC is supported for DAG.
- EVPN IRB interface cannot be a part of global Virtual Routing and Forwarding (VRF).
- Virtual Private LAN Service (VPLS) stitching is *not* supported.
- EVPN IRB is supported *only* on BDI interfaces.
- Remote MAC learned via EVPN-BGP cannot be controlled by MACsec/MAC limit features.
- Static MAC should *not* be configured as remote MAC address.
- As MAC scale is limited to 16K MACs at system level, remote MAC of 8000 scale is recommended.
- Remote MAC is learned at the rate of 400 PPS. Beyond this scale, you might encounter stale MACs or MAC stuck scenarios.
- A maximum number of 950 EVI is supported.
- EVPN IRB and MACsec features are *not* supported together.

## Configuring EVPN Integrated Routing and Bridging (L2 and L3 Anycast Gateway) and Data Center Interconnect or Border Leaf (Single Homing)

This section shows the configuration of EVPN IRB (L2 and L3 Anycast Gateway) and Data Center Interconnect (DCI) or Border Leaf (BL) Single Homing.

### Configure IP VRF

To configure IP Virtual Routing and Forwarding (VRF):

```
vrf definition evpn_vrf
rd 1.1.1.1:1
!
address-family ipv4
route-target export 1000:1
route-target import 1000:1
route-target export 1000:1 stitching
```

```

    route-target import 1000:1 stitching
  exit-address-family
!
address-family ipv6
  route-target export 1000:1
  route-target import 1000:1
  route-target export 1000:1 stitching
  route-target import 1000:1 stitching
exit-address-family

```

## Configure Layer 2 Virtual Private Network EVPN

To configure layer 2 virtual private network EVPN:

```

l2vpn evpn
  replication-type ingress
  router-id Loopback0
l2vpn evpn instance 10 vlan-based
l2vpn evpn instance 11 vlan-based
l2vpn evpn instance 12 vlan-based

```

## Configure Bridge Domain

To configure bridge domain:

```

bridge-domain 10
  member GigabitEthernet0/3/7 service-instance 10
  member evpn-instance 10
bridge-domain 11
  member GigabitEthernet0/3/7 service-instance 11
  member evpn-instance 11
bridge-domain 12
  member GigabitEthernet0/3/7 service-instance 12
  member evpn-instance 12

```

## Configure Bridge Domain IRB Interface

To configure bridge domain IRB interface:

```

interface BDI10
  mac-address 0011.1111.1111
  vrf forwarding evpn_vrf
  ip address 191.168.1.1 255.255.255.0
  ipv6 address 1968:1::1/64
end

interface BDI11
  mac-address 0011.1111.1112
  vrf forwarding evpn_vrf
  ip address 191.168.2.1 255.255.255.0
  ipv6 address 1969:2::1/64
end

```

## Configure BGP IRB

To configure BGP IRB:

```

router bgp 1000
  bgp router-id interface Loopback0

```

```

bgp log-neighbor-changes
neighbor 2.2.2.1 remote-as 1000
neighbor 2.2.2.1 ha-mode sso
neighbor 2.2.2.1 update-source Loopback0
!
address-family ipv4
  bgp additional-paths install
  network 1.1.1.1 mask 255.255.255.255
  neighbor 2.2.2.1 activate
  neighbor 2.2.2.1 send-community both
  neighbor 2.2.2.1 send-label
exit-address-family
!
address-family ipv6
  bgp additional-paths install
  no bgp recursion host
  neighbor 2.2.2.1 activate
  neighbor 2.2.2.1 send-community both
  neighbor 2.2.2.1 send-label
exit-address-family
!
address-family l2vpn evpn
  neighbor 2.2.2.1 activate
  neighbor 2.2.2.1 send-community both
exit-address-family
!
address-family ipv4 vrf evpn_vrf
  redistribute connected
exit-address-family
!
address-family ipv6 vrf evpn_vrf
  redistribute connected
exit-address-family

```

## Verification of EVPN Integrated Routing and Bridging (L2 and L3 Anycast Gateway) and Data Center Interconnect or Border Leaf (Single Homing) Configuration

This section shows the verification examples of EVPN IRB (L2 and L3 Anycast Gateway) and Data Center Interconnect or Border Leaf (Single Homing) configuration.

### Verify Device Tracking Database

Use **show device-tracking database** command to verify device tracking database:

```

Router#show device-tracking database
Binding Table has 16 entries, 7 dynamic
Codes: L - Local, S - Static, ND - Neighbor Discovery, ARP - Address Resolution Protocol,
DH4 - IPv4 DHCP, DH6 - IPv6 DHCP, PKT - Other Packet, API - API created
Preflevel flags (prlvl):
0001:MAC and LLA match      0002:Orig trunk          0004:Orig access
0008:Orig trusted trunk    0010:Orig trusted access  0020:DHCP assigned
0040:Cga authenticated     0080:Cert authenticated  0100:Statically assigned

```

Network Layer Address	Link Layer Address	Interface	bd	prlvl	age
state	Time left				
ARP 191.168.3.2	0000.0000.0003	Gi0/3/7	12	0005	174s



```

    REACHABLE 125 s
L   191.168.3.1                0011.1111.1113    BD12      12    0100  303mn
    REACHABLE
ARP 191.168.2.2                0000.0000.0002    Gi0/3/7   11    0005   3mn
    REACHABLE 117 s
L   191.168.2.1                0011.1111.1112    BD11      11    0100  303mn
    REACHABLE
ARP 191.168.1.2                0000.0000.0001    Gi0/3/7   10    0005  174s
    REACHABLE 131 s
L   191.168.1.1                0011.1111.1111    BD10      10    0100  303mn
    REACHABLE
L   FE80::211:11FF:FE11:1113    0011.1111.1113    BD12      12    0100  303mn
    REACHABLE
L   FE80::211:11FF:FE11:1112    0011.1111.1112    BD11      11    0100  303mn
    REACHABLE
L   FE80::211:11FF:FE11:1111    0011.1111.1111    BD10      10    0100  303mn
    REACHABLE
ND FE80::200:FF:FE00:3         0000.0000.0003    Gi0/3/7   12    0005  164s
    REACHABLE 146 s
ND FE80::200:FF:FE00:1         0000.0000.0001    Gi0/3/7   10    0005  164s
    REACHABLE 142 s
ND 1970:3::2                   0000.0000.0003    Gi0/3/7   12    0005  174s
    REACHABLE 126 s
L   1970:3::1                   0011.1111.1113    BD12      12    0100  303mn
    REACHABLE
L   1969:2::1                   0011.1111.1112    BD11      11    0100  303mn
    REACHABLE
ND 1968:1::2                   0000.0000.0001    Gi0/3/7   10    0005  174s
    REACHABLE 130 s
L   1968:1::1                   0011.1111.1111    BD10      10    0100  303mn
    REACHABLE

```

## Verify L2VPN EVPN Summary

Use **show l2vpn evpn summary** command to verify L2VPN EVPN IRB summary:

```

Router# show l2vpn evpn summary
L2VPN EVPN
  EVPN Instances (excluding point-to-point): 3
    VLAN Aware: 0
    VLAN Based: 3
    VLAN Bundle: 0
  Bridge Domains: 3
  BGP: ASN 1000, address-family l2vpn evpn configured
  Router ID: 1.1.1.1
  Label Allocation Mode: Per-BD
  Global Replication Type: Ingress
  ARP/ND Flooding Suppression: Enabled
  MAC Duplication: seconds 180 limit 5
  MAC Addresses: 3
    Local: 3
    Remote: 0
  Duplicate: 0
  IP Duplication: seconds 180 limit 5
  IP Addresses: 7
    Local: 7
    Remote: 0
    Duplicate: 0
  Advertise Default Gateway: No
  Default Gateway Addresses: 0
    Local: 0
    Remote: 0
  Global IP Local Learn: Enabled

```

```

IP local learning limits
  IPv4: 4 addresses per-MAC
  IPv6: 12 addresses per-MAC
IP local learning timers
  Down:      10 minutes
  Poll:      1 minutes
  Reachable: 5 minutes
  Stale:     30 minutes

```

## Verify L2VPN EVPN EVI

Use the **show l2vpn evpn evi detail** command to verify L2VPN EVPN EVI:

```

Router#show l2vpn evpn evi 10 detail
EVPN instance:      10 (VLAN Based)
RD:                 1.1.1.1:10 (auto)
Import-RTs:         1000:10
Export-RTs:         1000:10
Per-EVI Label:      none
State:              Established
Replication Type:   Ingress (global)
Encapsulation:      mpls
IP Local Learn:     Enabled (global)
Adv. Def. Gateway:  Disabled (global)
Bridge Domain:      10
  Ethernet-Tag:     0
  BUM Label:        16
  Per-BD Label:     17
  BDI Label:        none
  State:            Established
  Flood Suppress:   Attached
  Access If:        BDI10
  VRF:              evpn_vrf
  IPv4 IRB:         Enabled
  IPv6 IRB:         Enabled
Pseudoports:
  GigabitEthernet0/3/7 service instance 10
    Routes: 1 MAC, 3 MAC/IP

```

## Verify Platform Software Infrastructure Punt Statistics

Use the **show platform software infrastructure punt statistics** command to verify the platform software infrastructure punt statistics:

```

Router#show platform soft infrastructure punt statistics
UEA Punt Statistics

```

Global drops : 0

Queue Name	Rx count	Drop count
SW FORWARDING Q	403038	0
ROUTING PROTOCOL Q	159	0
ICMP Q	0	0
HOST Q	400	0
ACL LOGGING Q	0	0
STP Q	0	0
L2 PROTOCOL Q	0	0
MCAST CONTROL Q	0	0
BROADCAST Q	0	0
REP Q	0	0

BGP LDP Q	0	0
CONTROL Q	0	0
IP MPLS TTL Q	0	0
DEFAULT MCAST Q	0	0
MCAST ROUTE DATA Q	0	0
MCAST MISMATCH Q	0	0
RPF FAIL Q	0	0
ROUTING THROTTLE Q	0	0
MCAST Q	0	0
MPLS OAM Q	0	0
IP MPLS MTU Q	0	0
PTP Q	0	0
LINUX ND Q	0	0
KEEPALIVE Q	9256	0
ESMC Q	0	0
FPGA BFD Q	0	0
FPGA CCM Q	0	0
FPGA CFE Q	0	0
L2PT DUP Q	0	0
TDM CTRL Q	0	0
ICMP UNREACHABLE Q	402918	0
SSFP Q	0	0
MIRROT Q	0	0

## Verify Platform Software Infrastructure Inject

Use the **show platform software infrastructure inject** command to verify the platform software infrastructure inject:

```
Router#show platform software infrastructure inject
Statistics for L3 injected packets:
 27819 total inject pak, 24 failed
 0 sent, 532 prerouted
 24 non-CEF capable, 334 non-unicast
 4965 IP, 20885 IPv6
 0 MPLS, 0 Non-IP Tunnel
 0 UDLR tunnel, 0 P2MP replicated mcast
 0 Non-IP Fastswitched over Tunnel, 1945 legacy pak path
 0 Other packet
 0 IP fragmented
 25318 normal, 0 nexthop
 532 adjacency, 0 feature
 0 undefined
 0 pak find no adj, 0 no adj-id
 4991 sb alloc, 25850 sb local
 0 p2mcast failed count 0 p2mcast enqueue fail
 0 unicast dhc
 0 mobile ip
 105 IPv6 NA
 101 IPv6 NS
 0 Transport failed cases
 0 Grow packet buffer
 2 Cant-l3-inject-pkts
per feature packet inject statistics
 0 Feature multicast
 0 Feature Edge Switching Service
 0 Feature Session Border Controller
 0 Feature interrupt level
 0 Feature use outbound interface
 0 Feature interrupt level with OCE
 0 Feature ICMPv6 error message
 0 Feature Session Border Controller media packet injection
 0 Feature Tunnel Ethernet over GRE
```

```

0 Feature Secure Socket Layer Virtual Private Network
0 Feature EPC Wireshark injecting packets
0 Feature multicast overlay replication
Statistics for L2 injected packets:
356 total L2 inject pak, 0 failed
3 total BD inject pak, 0 failed
353 total EFP inject pak, 0 failed
0 total VLAN inject pak, 0 failed

```

## Verify BGP L2VPN EVPN Detail

Use the **show bgp l2vpn evpn detail** to verify BGP L2VPN EVPN detail:

```
Router#show bgp l2vpn evpn detail
```

```

Route Distinguisher: 1.1.1.1:10
BGP routing table entry for [2][1.1.1.1:10][0][48][000000000001][0][*]/20, version 1027
  Paths: (1 available, best #1, table evi_10)
  Not advertised to any peer
  Refresh Epoch 1
  Local
    :: (via default) from 0.0.0.0 (1.1.1.1)
      Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
      EVPN ESI: 00000000000000000000, Label1 17
      Extended Community: RT:1000:10
      rx pathid: 0, tx pathid: 0x0
      Updated on Dec 16 2020 22:39:43 IST
BGP routing table entry for [2][1.1.1.1:10][0][48][000000000001][32][191.168.1.2]/24, version 1021
  Paths: (1 available, best #1, table evi_10)
  Not advertised to any peer
  Refresh Epoch 1
  Local
    :: (via default) from 0.0.0.0 (1.1.1.1)
      Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
      EVPN ESI: 00000000000000000000, Label1 17, Label2 22
      Extended Community: RT:1000:1 RT:1000:10
      rx pathid: 0, tx pathid: 0x0
      Updated on Dec 16 2020 22:39:42 IST
BGP routing table entry for [2][1.1.1.1:10][0][48][000000000001][128][1968:1::2]/36, version 1022
  Paths: (1 available, best #1, table evi_10)
  Not advertised to any peer
  Refresh Epoch 1
  Local
    :: (via default) from 0.0.0.0 (1.1.1.1)
      Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
      EVPN ESI: 00000000000000000000, Label1 17, Label2 24
      Extended Community: RT:1000:1 RT:1000:10
      rx pathid: 0, tx pathid: 0x0
      Updated on Dec 16 2020 22:39:42 IST
BGP routing table entry for [2][1.1.1.1:10][0][48][000000000001][128][FE80::200:FF:FE00:1]/36, version 1023
  Paths: (1 available, best #1, table evi_10)
  Not advertised to any peer
  Refresh Epoch 1
  Local
    :: (via default) from 0.0.0.0 (1.1.1.1)
      Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
      EVPN ESI: 00000000000000000000, Label1 17, Label2 24
      Extended Community: RT:1000:1 RT:1000:10
      rx pathid: 0, tx pathid: 0x0
      Updated on Dec 16 2020 22:39:42 IST

```

```

Route Distinguisher: 1.1.1.1:11
BGP routing table entry for [2][1.1.1.1:11][0][48][000000000002][0][*]/20, version 1012
  Paths: (1 available, best #1, table evi_11)
  Not advertised to any peer
  Refresh Epoch 1
  Local
    :: (via default) from 0.0.0.0 (1.1.1.1)
      Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
      EVPN ESI: 00000000000000000000, Label1 19
      Extended Community: RT:1000:11
      rx pathid: 0, tx pathid: 0x0
      Updated on Dec 16 2020 22:39:13 IST
BGP routing table entry for [2][1.1.1.1:11][0][48][000000000002][32][191.168.2.2]/24, version 1011
  Paths: (1 available, best #1, table evi_11)
  Not advertised to any peer
  Refresh Epoch 1
  Local
    :: (via default) from 0.0.0.0 (1.1.1.1)
      Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
      EVPN ESI: 00000000000000000000, Label1 19, Label2 22
      Extended Community: RT:1000:1 RT:1000:11
      rx pathid: 0, tx pathid: 0x0
      Updated on Dec 16 2020 22:39:12 IST
Route Distinguisher: 1.1.1.1:12
BGP routing table entry for [2][1.1.1.1:12][0][48][000000000003][0][*]/20, version 1028
  Paths: (1 available, best #1, table evi_12)
  Not advertised to any peer
  Refresh Epoch 1
  Local
    :: (via default) from 0.0.0.0 (1.1.1.1)
      Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
      EVPN ESI: 00000000000000000000, Label1 21
      Extended Community: RT:1000:12
      rx pathid: 0, tx pathid: 0x0
      Updated on Dec 16 2020 22:39:43 IST
BGP routing table entry for [2][1.1.1.1:12][0][48][000000000003][32][191.168.3.2]/24, version 1024
  Paths: (1 available, best #1, table evi_12)
  Not advertised to any peer
  Refresh Epoch 1
  Local
    :: (via default) from 0.0.0.0 (1.1.1.1)
      Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
      EVPN ESI: 00000000000000000000, Label1 21, Label2 25
      Extended Community: RT:1000:2 RT:1000:12
      rx pathid: 0, tx pathid: 0x0
      Updated on Dec 16 2020 22:39:42 IST
BGP routing table entry for [2][1.1.1.1:12][0][48][000000000003][128][FE80::200:FF:FE00:3]/36, version 1026
  Paths: (1 available, best #1, table evi_12)
  Not advertised to any peer
  Refresh Epoch 1
  Local
    :: (via default) from 0.0.0.0 (1.1.1.1)
      Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
      EVPN ESI: 00000000000000000000, Label1 21, Label2 25
      Extended Community: RT:1000:2 RT:1000:12
      rx pathid: 0, tx pathid: 0x0
      Updated on Dec 16 2020 22:39:42 IST
Route Distinguisher: 1.1.1.1:10
BGP routing table entry for [3][1.1.1.1:10][0][32][1.1.1.1]/17, version 9
  Paths: (1 available, best #1, table evi_10)
  Not advertised to any peer

```

```

Refresh Epoch 1
Local
:: (via default) from 0.0.0.0 (1.1.1.1)
  Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
  Extended Community: RT:1000:10
  PMSI Attribute: Flags:0x0, Tunnel type:IR, length 4, label:16 tunnel identifier: 0000
0000
  rx pathid: 0, tx pathid: 0x0
  Updated on Dec 16 2020 17:33:58 IST

```

## EVPN-IRB DHCP v4 and v6 Relay over Segment Routing

**Table 22: Feature History**

Feature Name	Release Information	Description
EVPN-IRB DHCP v4 and v6 Relay over Segment Routing	Cisco IOS XE Bengaluru 17.6.1	<p>This feature introduces a specialised implementation of DHCP packets to support DHCPv4 and DHCPv6 in an EVPN Fabric with Distributed Anycast Gateways (DAGs) on the same Virtual Routing and Forwarding (VRF). It also avoids DHCP discovery packet floods across the fabric.</p> <p>The flooding suppression feature is also enhanced to intercept multicast or broadcast DHCP packets when DHCP relay is configured on the DAG to perform the required action and localize the scope of the service.</p> <p>This feature is not supported with Cisco ASR RSP3 module. It is only supported with Cisco ASR RSP2 module.</p> <p>This feature is only supported on NCS 4206 and NCS 4201/4202 routers.</p>

Prior to Cisco IOS XE Bengaluru Release 17.5.1, DHCP relay agent was not supported on EVPN fabric solution. EVPN IRB extends a link (layer 2 segment) across a routed backbone (fabric). As a result, multicast or broadcast packets reach all layer 2 segments that introduces problems like black holing etc and security breaches. In large EVPN deployments, the amount of multicast or broadcast traffic in the fabric can be overwhelming and can also overload DHCP server. In addition to scalability and performance impact, there is also a functional problem with DHCPv6 relay agents, which can lead to traffic drop and host reachability problems.

Starting with Cisco IOS XE Bengaluru Release 17.6.1, specialised handling of DHCP packets is implemented to support DHCPv4 and DHCPv6 in an EVPN Fabric with Distributed Anycast Gateways (DAGs) on the same Virtual Routing and Forwarding (VRF). It also avoids DHCP discovery packet floods across the fabric.

If the packets are flooded in the stretched layer 2, all the DHCP relay enabled DAGs relay the packet, which generates unnecessary workload on the DHCP servers. For certain specific scenarios, this also installs route on the wrong DAGs and causes outage.

To avoid the above-mentioned situation, First-Hop device handles the DHCP services and DHCP relay is enabled on IPv4 and IPv6 traffic for the same VRF.

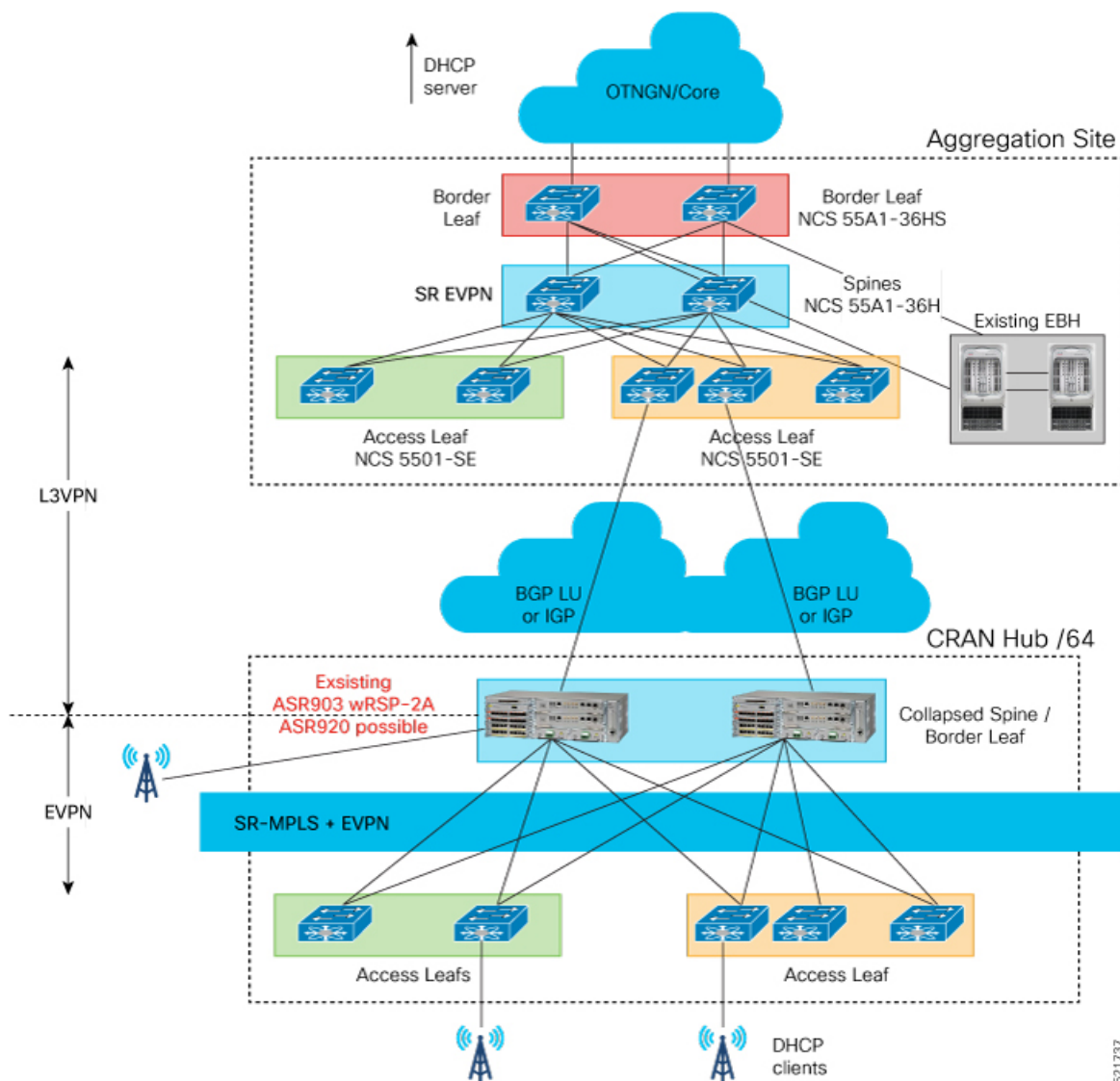
## How DHCP Relay Agent Works

DHCP relay agent is used to forward request and replies between DHCP server and client when the server is present on different link. The role of DHCP relay is to intercept the DHCP requests (broadcast or multicast) on a layer 2 network, convert it to layer 3 unicast traffic, and send to DHCP server. DHCP relay can be on any device in the network that has a layer 3 connectivity to the DHCP server. It is usually configured on the default gateway that faces the DHCP client. In an EVPN setup, the DHCP relay is configured on each DAG.

When the router receives multicast or broadcast DHCP packets from DHCP clients and if DHCP relay is configured on that interface, DHCP relay agent receives a copy of the DHCP packet and sends it to the DHCP server. There can be more than one relay or server on the same vlan for redundancy.

This feature ensures the DHCP packets only get flooded on local ports and not on the fabric ports. This stops other DHCP relay agents in the fabric to receive the DHCP packet. The following network topology is an example to show how DHCP relay agent works.

Figure 32: DHCP Relay Agent Network Topology



## Enhancement of Flooding Suppression Feature

Prior to Cisco IOS XE Bengaluru Release 17.6.1, the flooding suppression feature configured on DAGs in an IRB EVPN setup could suppress ARP and NP packets and minimize flooding of these packets into the EVPN fabric. Starting with Cisco IOS XE Bengaluru Release 17.6.1, this feature is enhanced to intercept multicast or broadcast DHCP packets when DHCP relay is configured on the DAG to perform the required action and localize the scope of the service.

Both DHCPv4 and DHCPv6 packets are suppressed across the fabric. All multicast or broadcast DHCP packets are forwarded unchanged in the local domain.



## Restrictions

- This feature is *only* supported on DAG deployments.
- This feature is *only* supported on Provider Edge (PE) routers.
- This feature cannot co-exist with DHCP snooping on the PE router.
- BD scale of 4095 is supported for EVPN IRB. However, only 100 of those BDs can be configured to use DHCP relay over EVPN IRB.
- When BDI is disabled or shut down, the DHCP packets remain suppressed. Thus, there is no flooding in the layer 2 domain.

## Configure DHCP Relay Flooding Suppression

To configure DHCP relay flooding suppression:

```
PE1(config)#service internal
PE1(config)#l2vpn evpn
PE1(config-evpn)#flooding-suppression
address-resolution Suppress flooding of Address Resolution and Neighbor
Discovery Protocol packets
PE1(config-evpn)#flooding-suppression dhcp-relay
disable Disable flooding suppression
PE1(config-evpn)#flooding-suppression dhcp-relay disable
```



### Note

- DHCP relay suppress flooding of DHCP packets to the fabric when DHCP relay is enabled locally.
- YANG model is not supported.
- The DHCP relay flooding suppression feature is enabled by default.

## Configure DHCP Relay on DAGs

**Global Configuration:**

```
ip dhcp relay information option vpn
```

**DHCP Server in the Same Tenant IP-VRF:**



### Note

Same tenant IP-VRFs are the VRFs that are part of the same leaf node.

```
!! Unique Loopback IPv4/IPv6 in the tenant IP-VRF
interface Loopback192
description DHCP-Relay Source Interface
vrf forwarding red
ip address 10.10.10.4 255.255.255.255
ipv6 address 2001:10::4/128
end

interface BDI12
```

```

mac-address 0012.0012.0012
vrf forwarding red
ip dhcp relay source-interface Loopback192
ip address 192.168.12.254 255.255.255.0
ip helper-address 10.10.10.10
ipv6 address 2001:12::254/64
ipv6 dhcp relay destination 2001:10::10
ipv6 dhcp relay source-interface Loopback192
end!

```

### DHCP Server in Different Tenant IP-VRF:



#### Note

Different tenant IP-VRFs are not part of the same leaf node and they need the IRB functionality.

```

!! Unique Loopback IPv4/IPv6 in the tenant IP-VRF
interface Loopback192
description DHCP-Relay Source Interface
vrf forwarding green
ip address 10.10.10.4 255.255.255.255
ipv6 address 2001:10::4/128
end
interface BDI12
mac-address 0012.0012.0012
vrf forwarding red
ip dhcp relay source-interface Loopback192
ip address 192.168.12.254 255.255.255.0
ip helper-address vrf green 10.10.10.10
ipv6 address 2001:12::254/64
ipv6 dhcp relay vrf green destination 2001:10::10
ipv6 dhcp relay source-interface Loopback192
end!

```

### DHCP Server in Global VRF:

```

!! No need for Unique Loopback IPv4/IPv6
interface BDI12
mac-address 0012.0012.0012
vrf forwarding red
ip address 192.168.12.254 255.255.255.0
ip helper-address global 10.10.10.10
ipv6 address 2001:12::254/64
ipv6 dhcp relay global destination 2001:10::10
end!

```

## Verification of DHCPv4 and DHCPv6 Relay Configuration

Use the **show device-tracking policies** command to verify all the SISF feature policies attached to bridge domain.

```

PE3#show device-tracking policies
Target      Type Policy                               Feature      Target range
bd 11       bd  evpn-no-device-track Device-tracking bd all
bd 11       bd  evpn-flood-suppress  Flooding Suppress bd all

```

Use the **show flooding-suppression policy** command to verify the settings of a policy.

```

PE3#show flooding-suppression policy
PE3#show flooding-suppression policy evpn-flood-suppress
Flooding suppress policy evpn-flood-suppress configuration:
Suppressing NDP

```

```

Suppressing DHCPv6
Suppressing ARP
Suppressing DHCPv4
mode:No-Proxy multicast resolution requests
Policy evpn-flood-suppress is applied on the following targets:
Target          Type Policy          Feature          Target range
bd 11           BD  evpn-flood-suppress Flooding Suppr bd all
bd 12           BD  evpn-flood-suppress Flooding Suppr bd all

```

Use the **show device-tracking counters** *bridge-domain-id* to verify the counters.

```

PE-1#show device-tracking counters bd 11
Received messages on bd 11 :
Protocol          Protocol message
NDP               RS[4] RA[4] NS[1777] NA[2685]
DHCPv6
ARP              REQ[12] REP[1012]
DHCPv4
ACD&DAD          --[8]
:
:
Limited Broadcast to Local message on bd 11 :
Type              Protocol message
NDP
DHCPv6            SOL[1] REQ[1] REB[1]
ARP
DHCPv4            DIS[1] REQ4[1]

```

Use the **show l2fib output-list** and **show l2fib bridge-domain** *bd -id table unicast* commands to verify the information about Layer 2 Forwarding Information Base (L2FIB).

```

PE3#show l2fib output-list
ID BD Port Flags
-----
1035 11 1 local port list
1036 12 0 local port list
5120 1 0 flood list
5130 11 4 flood list
5131 12 1 flood list
PE3#show l2fib output-list 1035
ID : 1035
Bridge Domain : 11
Reference Count : 3
Flags : local port list
Port Count : 1
Port(s) : BD_PORT Gi0/3/2:11
PE3#show l2fib bridge-domain 11 table unicast
MAC Address Adjacency
-----
7069.5a39.ef8a BD_PORT Gi0/3/2:11
7069.5a39.ef94 MPLS_UC PL:1(1) T:MPLS_UC [MAC]17@99.99.99.1
7486.0bc4.d4d4 BD_PORT Gi0/3/2:11
ffff.ffff.ffff Olist: 1035, Ports: 1

```

Use the **show platform software l2fib f0 bdbdomain idunicast all** command to verify the global bridge domain table for MAC and layer 2 multicast.

```

PE-1#show platform software l2fib f0 bd 10 unicast all
MAC          BD          Nhop type          Nhop Idx          Flags
-----
ffff.ffff.ffff          10          olist          1034
static

```

Use the **show platform software l2fib f0 mlist index** *Nhop Idx* command to verify the output list of the global bridge domain table for MAC and layer 2 multicast.

```
PE-1#show plat soft l2fib f0 mlist index 1034
L2FIB Mlist entries
```

```
Type Index AOM ID CPP Info
```

```
efp 408011 aom id: 149, CPP info: 0x1808bcc (created)
```

Use the **show platform software dpidb ethernet efp interface** *interface-name* command to verify <need information>

```
PE-1#sh platform software dpidb ethernet efp interface gigabitEthernet 0/0/0
DPIDB for interface GigabitEthernet0/0/0 (Ethernet Flow Points)
EFP ID: 10, dpidb index: 0x408011
```

Use the **show platform software dpidb ethernet efp interface show platform hardware pp active bridge-domain id** *domain id* command to verify the view the Ethernet Flow Point (EFP) information.

```
PE-1#show platform hard pp ac bridge-domain id 10
Bridge Domain Details
:
Nile Vlan Compression Table
```

type	brdgeD_index	fid_index	floodtype	Entry MET
lp_access	NA	24877	3	184341

Use the **show platform software infrastructure inject** command to verify the platform software infrastructure inject:

```
PE-1#show platform software infrastructure inject
:
Statistics for L2 injected packets:
1 total L2 inject pak, 0 failed
0 total BD inject pak, 0 failed
1 total BD-local inject pak, 0 failed
0 total EFP inject pak, 0 failed
0 total VLAN inject pak, 0 failed
```

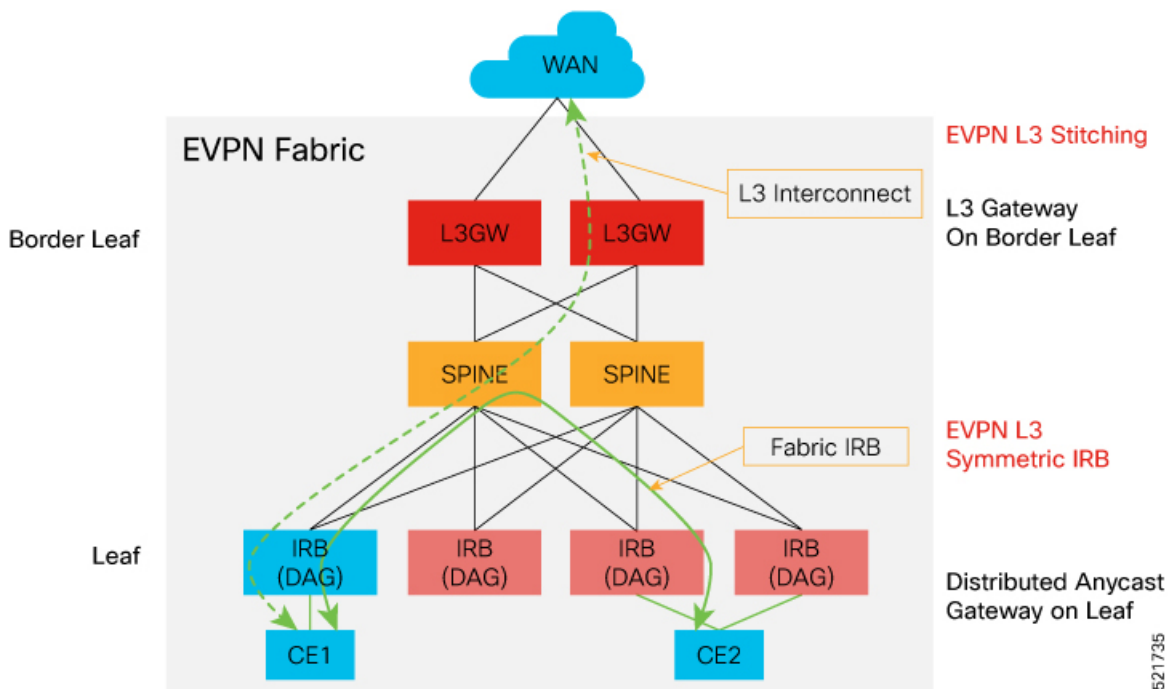
# Stitching of Subnet Route from EVPN to L3VPN

**Table 23: Feature History**

Feature Name	Release	Description
Stitching of Subnet Route from EVPN to L3VPN	Cisco IOS XE Bengaluru 17.6.1	<p>This feature introduces the collapsed spine and border leaf node in the network topology of single homing DAGs with symmetric IRB, inter-subnet layer 3 traffic within fabric and inter-subnet layer 3 stitching through layer 3 border gateway. The hosts participating in fabric IRB are directly attached with the collapsed spine and border leaf node.</p> <p>This feature is not supported with Cisco ASR RSP3 module. It is only supported with Cisco ASR RSP2 module.</p> <p>This is only supported on NCS 4206 and NCS 4201/4202 routers.</p>

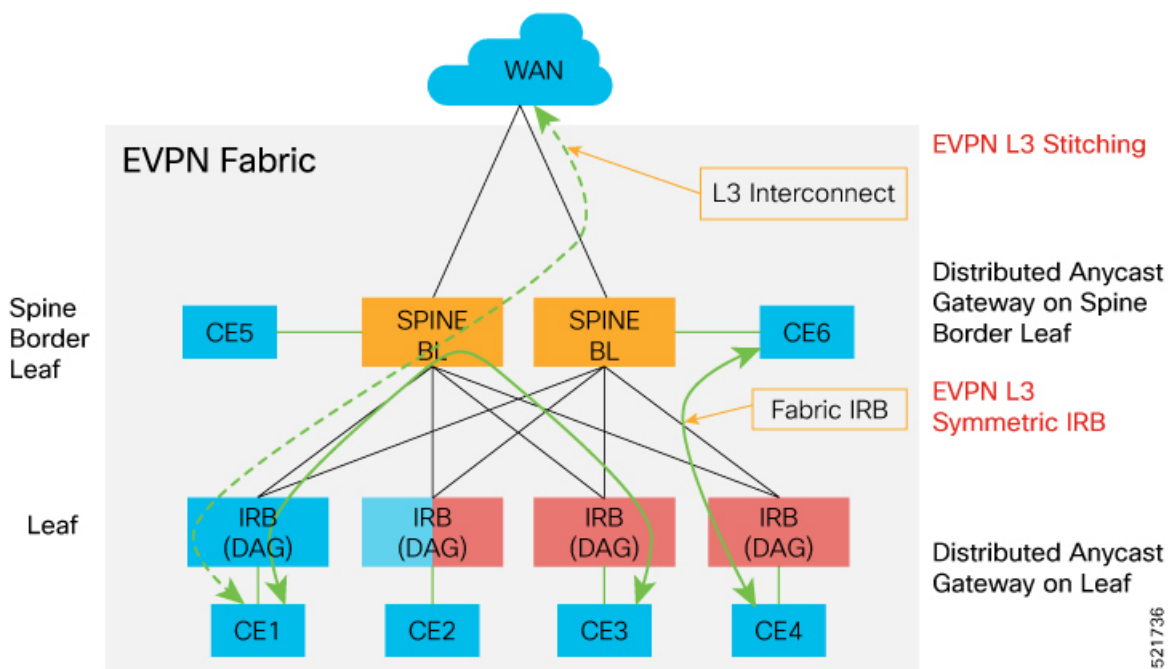
Prior to Cisco IOS XE Bengaluru Release 17.6.1, the hosts participating in fabric IRB were not attached directly with the spine and border leaf nodes. As a result, the hosts had to pass through the spine and border leaf nodes to reach the network. The following figure shows the network topology that illustrates single homing DAGs with symmetric IRB, inter-subnet layer 3 traffic within fabric and inter-subnet layer 3 stitching through layer 3 border gateway.

Figure 33: Single Homing DAGs with Symmetric IRB Network Topology



Starting with Cisco IOS XE Bengaluru Release 17.6.1, the hosts participating in fabric IRB are directly attached with the collapsed spine and border leaf node. The following figure shows the network topology that illustrates single homing DAGs with symmetric IRB, inter-subnet layer 3 traffic within fabric and inter-subnet layer 3 stitching through spine border gateway.

Figure 34: Collapsed Spine or Leaf or RR on a Single Node Network Topology



**Leaf Node:** The Provider Edge (PE) node that provides EVPN layer 2 or layer 3 services for Customer Edge (CE) devices (wireless nodes). DAG functionality is enabled for layer 3 services. Each PE may host a management VLAN for all the devices that are co-located with the PE. This VLAN is reachable via the management DAG from aggregation or core layers. This DAG is part of a management VRF. The in-band management interface of PE (statically configured loopback interface) is also a part of the same VRF. DHCPv4 and DHCPv6 relay functionality with centralized DHCP server are located beyond the aggregation domain. DHCP server is a part of the same VRF with the corresponding IRB interfaces.

**Spine Node:** Performs SR and MPLS forwarding among leaf nodes within a single CRAN hub (RR in EVPN address family).

**Border Leaf Node:** Forwards traffic between CRAN Hub and the aggregation layer. This node stitches EVPN Layer 3 routes and L3VPN/6VPE BGP routes.

## Restrictions

- DHCP Snooping and EVPN IRB snooping cannot be enabled on the same bridge domain.
- Stitching on a collapsed Border Leaf (BL) or Spine or Leaf node is *not* supported.
- EVPN IRB feature and Hot Standby Router Protocol (HSRP) or Virtual Router Redundancy Protocol (VRRP) feature are *not* supported simultaneously on a router.
- EVPN Layer 2 is *not* supported on Media Access Control Security (MACsec).
- Only Virtual MAC is supported for DAG.
- Asymmetric IRB and centralized IRB are *not* supported.
- EVPN IRB interface cannot be a part of global Virtual Routing and Forwarding (VRF).
- Virtual Private LAN Service (VPLS) stitching is *not* supported.
- EVPN IRB is supported *only* on BDI interfaces.
- Remote MAC learned via EVPN-BGP cannot be controlled by MACsec/MAC limit features.
- Static MAC should not be configured as remote MAC address.
- As MAC scale is limited to 16K MACs at system level, remote MAC of 8000 scale is recommended.
- Remote MAC is learned at the rate of 400 PPS. Beyond this scale, you might encounter stale MACs or MAC stuck scenarios.
- A maximum number of 950 EVI is supported.
- EVPN IRB and MACsec features are *not* supported together.
- Only 100 BDs can be configured to use DHCP relay over EVPN IRB.

## Scale

The following table shows the scale for Cisco ASR RSP2 module.

Table 24: Scale for Cisco ASR RSP2 Module

EVI	BD	MAC per BD	MACs Per System (Local and Remote)	Scale (Default Template)
950	4095  <b>Note</b> BD scale of 4095 is supported for EVPN IRB. However, only 100 of those BDs can be configured to use DHCP relay over EVPN IRB.	16K	16K (A maximum of <= 8K remote MAC is recommended) MACs Per system (local and remote)	IPV4 ROUTES = 20000  IPV6 ROUTES = 4000  Maximum VRF = 128  Maximum BDI = 4095  255 VMAC

## Configure BGP L3VPN Subnet Advertisement (Stitching) for Collapsed Spine or Leaf or RR on a Single Node

To configure BGP L3VPN subnet advertisement (stitching) for collapsed spine or leaf or RR on a single node:

```

vrf definition evpn_vrf
rd 1.1.1.1:1
!
address-family ipv4
route-target export 1000:1
route-target import 1000:1
route-target export 1000:1 stitching
route-target import 1000:1 stitching
exit-address-family
!
address-family ipv6
route-target export 1000:1
route-target import 1000:1
route-target export 1000:1 stitching
route-target import 1000:1 stitching
exit-address-family
l2vpn evpn
replication-type ingress
router-id Loopback0
l2vpn evpn instance 10 vlan-based
l2vpn evpn instance 11 vlan-based
l2vpn evpn instance 12 vlan-based

```



## Verification of BGP L3VPN Subnet Advertisement (Stitching) for Collapsed Spine or Leaf or RR on a Single Node Configuration

Use the following commands to verify BGP L3VPN subnet advertisement (stitching) for collapsed spine or leaf or RR on a single node configuration.

Use the **show device-tracking database** command to verify device tracking database:

```
PE1#show device-tracking database
Binding Table has 16 entries, 7 dynamic
Codes: L - Local, S - Static, ND - Neighbor Discovery, ARP - Address Resolution Protocol,
DH4 - IPv4 DHCP, DH6 - IPv6
DHCP, PKT - Other Packet, API - API created
Preflevel flags (prlvl):
0001:MAC and LLA match 0002:Orig trunk 0004:Orig access
0008:Orig trusted trunk 0010:Orig trusted access 0020:DHCP assigned
0040:Cga authenticated 0080:Cert authenticated 0100:Statically assigned
Network Layer Address Link Layer Address Interface bd prlvl age state Time left
ARP 191.168.3.2 0000.0000.0003 Gi0/3/7 12 0005 174s REACHABLE 125 s
L 191.168.3.1 0011.1111.1113 BD12 12 0100 303mn REACHABLE
ARP 191.168.2.2 0000.0000.0002 Gi0/3/7 11 0005 3mn REACHABLE 117 s
L 191.168.2.1 0011.1111.1112 BD11 11 0100 303mn REACHABLE
ARP 191.168.1.2 0000.0000.0001 Gi0/3/7 10 0005 174s REACHABLE 131 s
L 191.168.1.1 0011.1111.1111 BD10 10 0100 303mn REACHABLE
L FE80::211:11FF:FE11:1113 0011.1111.1113 BD12 12 0100 303mn REACHABLE
L FE80::211:11FF:FE11:1112 0011.1111.1112 BD11 11 0100 303mn REACHABLE
L FE80::211:11FF:FE11:1111 0011.1111.1111 BD10 10 0100 303mn REACHABLE
ND FE80::200:FF:FE00:3 0000.0000.0003 Gi0/3/7 12 0005 164s REACHABLE 146 s
ND FE80::200:FF:FE00:1 0000.0000.0001 Gi0/3/7 10 0005 164s REACHABLE 142 s
ND 1970:3::2 0000.0000.0003 Gi0/3/7 12 0005 174s REACHABLE 126 s
L 1970:3::1 0011.1111.1113 BD12 12 0100 303mn REACHABLE
L 1969:2::1 0011.1111.1112 BD11 11 0100 303mn REACHABLE
ND 1968:1::2 0000.0000.0001 Gi0/3/7 10 0005 174s REACHABLE 130 s
L 1968:1::1 0011.1111.1111 BD10 10 0100 303mn REACHABLE
```

Use the **show l2vpn evpn summary** command to verify L2VPN EVPN IRB summary:

```
PE1# show l2vpn evpn summary
L2VPN EVPN
EVPN Instances (excluding point-to-point): 3
VLAN Aware: 0
VLAN Based: 3
VLAN Bundle: 0
Bridge Domains: 3
BGP: ASN 1000, address-family l2vpn evpn configured
Router ID: 1.1.1.1
Label Allocation Mode: Per-BD
Global Replication Type: Ingress
ARP/ND Flooding Suppression: Enabled
MAC Duplication: seconds 180 limit 5
MAC Addresses: 3
Local: 3
Remote: 0
Duplicate: 0
IP Duplication: seconds 180 limit 5
IP Addresses: 7
Local: 7
Remote: 0
Duplicate: 0
Advertise Default Gateway: No
Default Gateway Addresses: 0
Local: 0
Remote: 0
```

```

Global IP Local Learn: Enabled
IP local learning limits
IPv4: 4 addresses per-MAC
IPv6: 12 addresses per-MAC
IP local learning timers
Down: 10 minutes
Poll: 1 minutes
Reachable: 5 minutes
Stale: 30 minutes

```

Use the **show l2vpn evpn evi detail** command to verify L2VPN EVPN EVI:

```

PE1#show l2vpn evpn evi 10 detail
EVPN instance: 10 (VLAN Based)
RD: 1.1.1.1:10 (auto)
Import-RTs: 1000:10
Export-RTs: 1000:10
Per-EVI Label: none
State: Established
Replication Type: Ingress (global)
Encapsulation: mpls
IP Local Learn: Enabled (global)
Adv. Def. Gateway: Disabled (global)
Bridge Domain: 10
Ethernet-Tag: 0
BUM Label: 16
Per-BD Label: 17
BDI Label: none
State: Established
Flood Suppress: Attached
Access If: BDI10
VRF: evpn_vrf
IPv4 IRB: Enabled
IPv6 IRB: Enabled
Pseudoports:
GigabitEthernet0/3/7 service instance 10
Routes: 1 MAC, 3 MAC/IP

```

Use the **show platform software infrastructure punt statistics** command to verify the platform software infrastructure punt statistics:

```

PE1#show platform software infrastructure punt statistics
UEA Punt Statistics
Global drops : 0
Queue Name | Rx count | Drop count
-----+-----+-----
SW FORWARDING Q | 403038 | 0
ROUTING PROTOCOL Q | 159 | 0
ICMP Q | 0 | 0
HOST Q | 400 | 0
ACL LOGGING Q | 0 | 0
STP Q | 0 | 0
L2 PROTOCOL Q | 0 | 0
MCAST CONTROL Q | 0 | 0
BROADCAST Q | 0 | 0
REP Q | 0 | 0
BGP LDP Q | 0 | 0
CONTROL Q | 0 | 0
IP MPLS TTL Q | 0 | 0
DEFAULT MCAST Q | 0 | 0
MCAST ROUTE DATA Q | 0 | 0
MCAST MISMATCH Q | 0 | 0
RPF FAIL Q | 0 | 0
ROUTING THROTTLE Q | 0 | 0
MCAST Q | 0 | 0

```

```

MPLS OAM Q | 0 | 0
IP MPLS MTU Q | 0 | 0
PTP Q | 0 | 0
LINUX ND Q | 0 | 0
KEEPALIVE Q | 9256 | 0
ESMC Q | 0 | 0
FPGA BFD Q | 0 | 0
FPGA CCM Q | 0 | 0
FPGA CFE Q | 0 | 0
L2PT DUP Q | 0 | 0
TDM CTRL Q | 0 | 0
ICMP UNREACHABLE Q | 402918 | 0
SSFP Q | 0 | 0
MIRROT Q | 0 | 0

```

Use the **show platform software infrastructure inject** command to verify the platform software infrastructure inject:

```

PE1#show platform software infrastructure inject
Statistics for L3 injected packets:
27819 total inject pak, 24 failed
0 sent, 532 prerouted
24 non-CEF capable, 334 non-unicast
4965 IP, 20885 IPv6
0 MPLS, 0 Non-IP Tunnel
0 UDLR tunnel, 0 P2MP replicated mcast
0 Non-IP Fastswitched over Tunnel, 1945 legacy pak path
0 Other packet
0 IP fragmented
25318 normal, 0 nexthop
532 adjacency, 0 feature
0 undefined
0 pak find no adj, 0 no adj-id
4991 sb alloc, 25850 sb local
0 p2mcast failed count 0 p2mcast enqueue fail
0 unicast dhc
0 mobile ip
105 IPv6 NA
101 IPv6 NS
0 Transport failed cases
0 Grow packet buffer
2 Cant-l3-inject-pkts
per feature packet inject statistics
0 Feature multicast
0 Feature Edge Switching Service
0 Feature Session Border Controller
0 Feature interrupt level
0 Feature use outbound interface
0 Feature interrupt level with OCE
0 Feature ICMPv6 error message
0 Feature Session Border Controller media packet injection
0 Feature Tunnel Ethernet over GRE
0 Feature Secure Socket Layer Virtual Private Network
0 Feature EPC Wireshark injecting packets
0 Feature multicast overlay replication
Statistics for L2 injected packets:
356 total L2 inject pak, 0 failed
3 total BD inject pak, 0 failed
353 total EFP inject pak, 0 failed
0 total VLAN inject pak, 0 failed

```

Use the **show bgp l2vpn evpn detail** to verify BGP L2VPN EVPN detail:

```

PE1#show bgp l2vpn evpn detail
Route Distinguisher: 1.1.1.1:10

```

```

BGP routing table entry for [2][1.1.1.1:10][0][48][000000000001][0][*]/20, version 1027
Paths: (1 available, best #1, table evi_10)
Not advertised to any peer
Refresh Epoch 1
Local
:: (via default) from 0.0.0.0 (1.1.1.1)
Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
EVPN ESI: 00000000000000000000, Label1 17
Extended Community: RT:1000:10
rx pathid: 0, tx pathid: 0x0
Updated on Dec 16 2020 22:39:43 IST
BGP routing table entry for [2][1.1.1.1:10][0][48][000000000001][32][191.168.1.2]/24, version
1021
Paths: (1 available, best #1, table evi_10)
Not advertised to any peer
Refresh Epoch 1
Local
:: (via default) from 0.0.0.0 (1.1.1.1)
Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
EVPN ESI: 00000000000000000000, Label1 17, Label2 22
Extended Community: RT:1000:1 RT:1000:10
rx pathid: 0, tx pathid: 0x0
Updated on Dec 16 2020 22:39:42 IST
BGP routing table entry for [2][1.1.1.1:10][0][48][000000000001][128][1968:1::2]/36, version
1022
Paths: (1 available, best #1, table evi_10)
Not advertised to any peer
Refresh Epoch 1
Local
:: (via default) from 0.0.0.0 (1.1.1.1)
Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
EVPN ESI: 00000000000000000000, Label1 17, Label2 24
Extended Community: RT:1000:1 RT:1000:10
rx pathid: 0, tx pathid: 0x0
Updated on Dec 16 2020 22:39:42 IST
BGP routing table entry for [2][1.1.1.1:10][0][48][000000000001][128][FE80::200:FF:FE00:1]/36,
version 1023
Paths: (1 available, best #1, table evi_10)
Not advertised to any peer
Refresh Epoch 1
Local
:: (via default) from 0.0.0.0 (1.1.1.1)
Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
EVPN ESI: 00000000000000000000, Label1 17, Label2 24
Extended Community: RT:1000:1 RT:1000:10
rx pathid: 0, tx pathid: 0x0
Updated on Dec 16 2020 22:39:42 IST
Route Distinguisher: 1.1.1.1:11
BGP routing table entry for [2][1.1.1.1:11][0][48][000000000002][0][*]/20, version 1012
Paths: (1 available, best #1, table evi_11)
Not advertised to any peer
Refresh Epoch 1
Local
:: (via default) from 0.0.0.0 (1.1.1.1)
Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
EVPN ESI: 00000000000000000000, Label1 19
Extended Community: RT:1000:11
rx pathid: 0, tx pathid: 0x0
Updated on Dec 16 2020 22:39:13 IST
BGP routing table entry for [2][1.1.1.1:11][0][48][000000000002][32][191.168.2.2]/24, version
1011
Paths: (1 available, best #1, table evi_11)
Not advertised to any peer
Refresh Epoch 1

```

```

Local
:: (via default) from 0.0.0.0 (1.1.1.1)
Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
EVPN ESI: 00000000000000000000, Label1 19, Label2 22
Extended Community: RT:1000:1 RT:1000:11
rx pathid: 0, tx pathid: 0x0
Updated on Dec 16 2020 22:39:12 IST
Route Distinguisher: 1.1.1.1:12
BGP routing table entry for [2][1.1.1.1:12][0][48][000000000003][0][*]/20, version 1028
Paths: (1 available, best #1, table evi_12)
Not advertised to any peer
Refresh Epoch 1
Local
:: (via default) from 0.0.0.0 (1.1.1.1)
Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
EVPN ESI: 00000000000000000000, Label1 21
Extended Community: RT:1000:12
rx pathid: 0, tx pathid: 0x0
Updated on Dec 16 2020 22:39:43 IST
BGP routing table entry for [2][1.1.1.1:12][0][48][000000000003][32][191.168.3.2]/24, version
1024
Paths: (1 available, best #1, table evi_12)
Not advertised to any peer
Refresh Epoch 1
Local
:: (via default) from 0.0.0.0 (1.1.1.1)
Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
EVPN ESI: 00000000000000000000, Label1 21, Label2 25
Extended Community: RT:1000:2 RT:1000:12
rx pathid: 0, tx pathid: 0x0
Updated on Dec 16 2020 22:39:42 IST
BGP routing table entry for [2][1.1.1.1:12][0][48][000000000003][128][FE80::200:FF:FE00:3]/36,
version 1026
Paths: (1 available, best #1, table evi_12)
Not advertised to any peer
Refresh Epoch 1
Local
:: (via default) from 0.0.0.0 (1.1.1.1)
Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
EVPN ESI: 00000000000000000000, Label1 21, Label2 25
Extended Community: RT:1000:2 RT:1000:12
rx pathid: 0, tx pathid: 0x0
Updated on Dec 16 2020 22:39:42 IST
Route Distinguisher: 1.1.1.1:10
BGP routing table entry for [3][1.1.1.1:10][0][32][1.1.1.1]/17, version 9
Paths: (1 available, best #1, table evi_10)
Not advertised to any peer
Refresh Epoch 1
Local
:: (via default) from 0.0.0.0 (1.1.1.1)
Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
Extended Community: RT:1000:10
PMSI Attribute: Flags:0x0, Tunnel type:IR, length 4, label:16 tunnel identifier: 0000 0000
rx pathid: 0, tx pathid: 0x0
Updated on Dec 16 2020 17:33:58 IST

```

