Segment Routing Configuration Guide, Cisco IOS XE 16 (Cisco ASR 900 Series)

First Published: 2017-12-07

Last Modified: 2020-04-09

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.

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 NONINFRINGEMENT 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.

CONTENTS

CHAPTER 1

Feature History

The following table lists the new and modified features supported in the MPLS Layer 2 VPNs Configuration Guide in Cisco IOS XE 16 releases.

Feature Name	Cisco IOS XE Release
Segment Routing uLoop Avoidance	16.12.1a
Segment Routed Traffic Engineering On-Demand Next Hop Color-extended Community Support and Constraints	16.11.1a
Segment Routing Label Recirculation for TI-LFA	
Seven Segment Routing-Traffic Engineering Label Support	
Segment Routing – Traffic Engineering (SR-TE)	16.6.1
Segment Routing Enhancement	

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?

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:

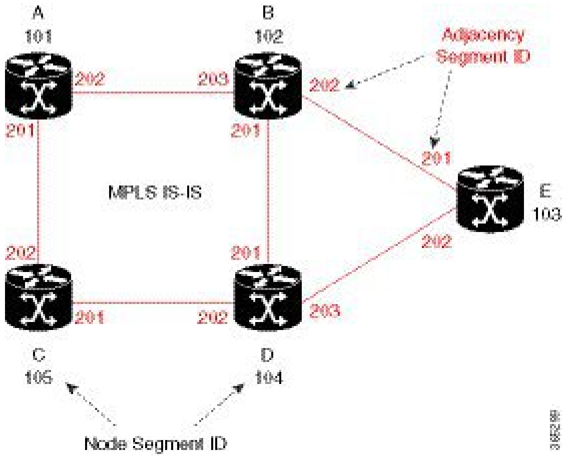
- Prefix SID—A segment ID that contains an IP address prefix calculated by an IGP in the service provider core network. Prefix SIDs are globally unique. A node SID is a special form of prefix SID that contains the loop-back address of the node as the prefix. It is advertised as an index into the node-specific 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

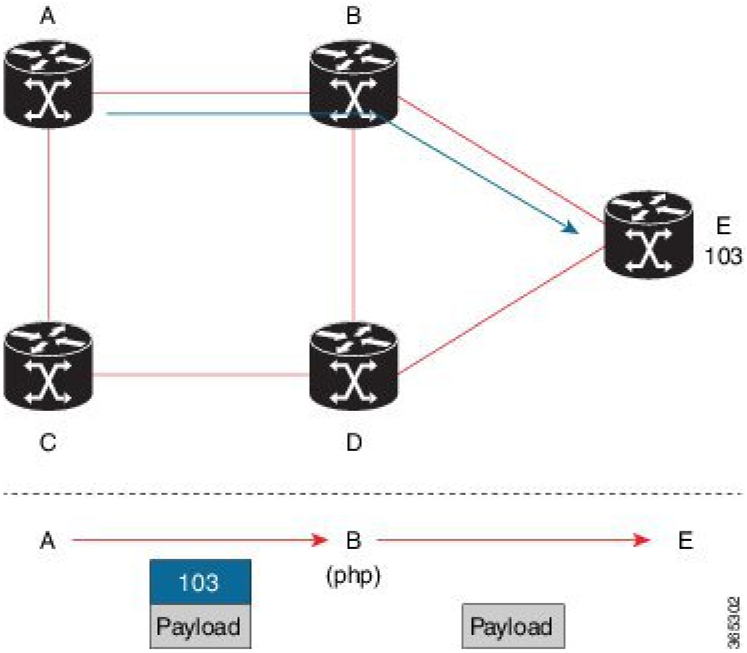
The following figure illustrates an MPLS network with five routers using Segment Routing, IS-IS, a label range of 100 to 199 for node IDs, and 200 and higher for adjacency IDs. IS-IS would distribute IP prefix reachability alongside segment ID (the MPLS label) across the network.

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



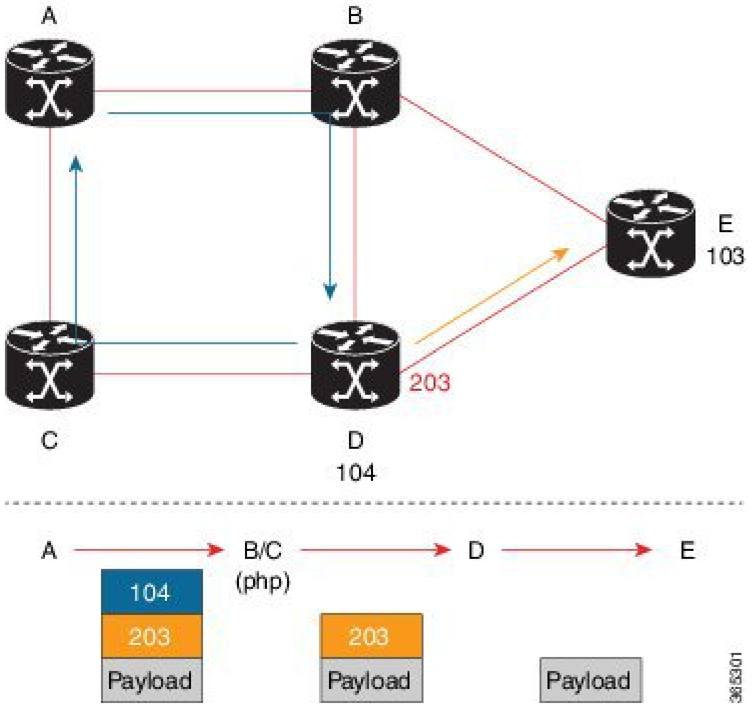
In the previous example, any router sending traffic to router E would push label 103 (router E node segment identifier) to forward traffic using the IS-IS shortest path. The MPLS label-swapping operation at each hop preserves label 103 until the packet arrives at E (Figure 2). On the other hand, adjacency 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.

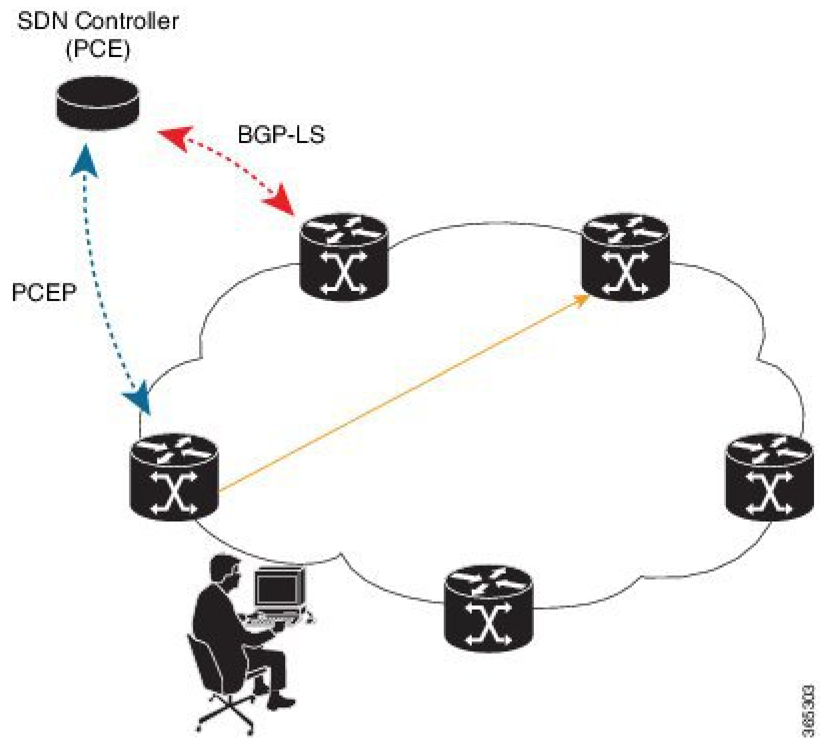
Figure 3: Destination Path of Router E



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."

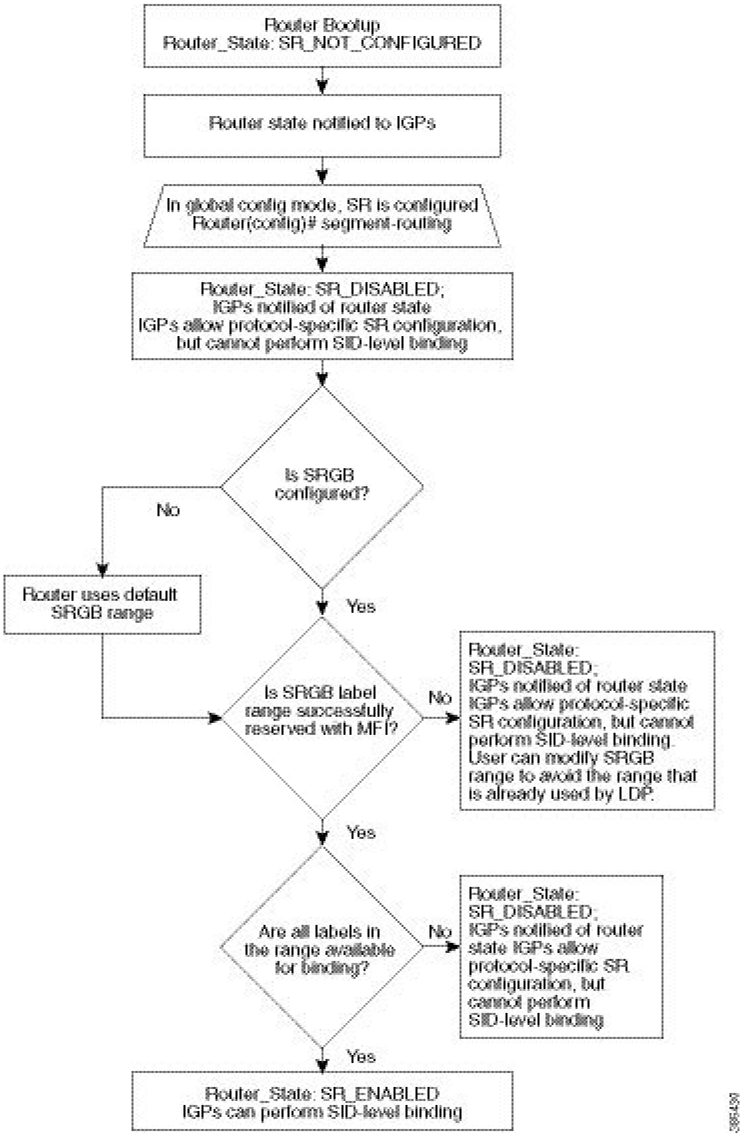
- Centralized Traffic Engineering—
  - Controllers and orchestration platforms can interact with Segment Routing traffic engineering for centralized optimization, such as WAN optimization.
  - Network changes such as congestion can trigger an application to optimize (recompute) the placement of segment 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


Router Bootup
Router_State: SR_NOT_CONFIGURED
Router state notified to IGPs
In global config mode, SR is configured
Router(config)# segment-routing
Router_State: SR_DISABLED;
IGPs notified of router state
IGPs allow protocol-specific SR configuration,
but cannot perform SID-level binding
Is SRGB configured?
No
Router uses default SRGB range
Yes
Is SRGB label range successfully reserved with MFI?
No
Router_State: SR_DISABLED;
IGPs notified of router state
IGPs allow protocol-specific SR configuration, but cannot perform SID-level binding.
User can modify SRGB range to avoid the range that is already used by LDP.
Yes
Are all labels in the range available for binding?
No
Router_State: SR_DISABLED;
IGPs notified of router state IGPs allow protocol-specific SR configuration, but cannot perform SID-level binding
Yes
Router_State: SR_ENABLED
IGPs can perform SID-level binding

- Segment Routing must be globally enabled on the chassis *before* enabling it on the IGPs, 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

- Segment routing is supported on the Cisco ASR 900 with RSP2 and RSP3 modules.
- The Cisco ASR 900 router with RSP2 module supports five label stacks. The Cisco ASR 900 router with RSP3 module supports four label stacks.
- 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.
- While configuring FRR, *carrier-delay down msec 1* 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 43.

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 configures in the global SR mode.

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

Disabling the Mapping Server

When computing SIDs for prefixes, IGPs consider 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

- 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 9 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* + 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 IGPs export 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 IGPs communicate 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

- 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 IGPs use 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

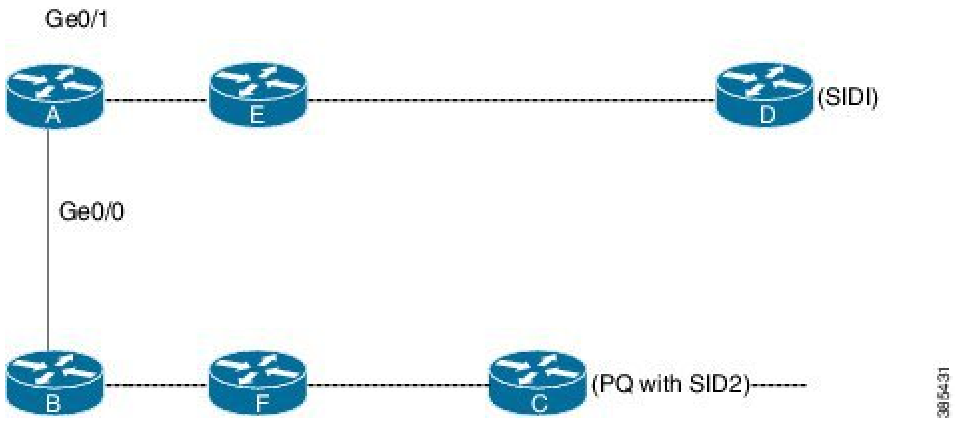
IS-IS announces the SR information to TE when it detects that both, IS-IS SR and TE are enabled for at least one level. IS-IS announce only the information that is obtained from the level for which TE is configured.

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-TILFA.
- In Cisco IOS XE Release 3.18S, 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.
- In Cisco IOS XE Release 3.18S, four MPLS label push is supported. TI-LFA tunnel carries a maximum of two labels and the other two labels are for services.

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-Beakers

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-srmpls)# 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-srmpls)#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.

```
Roouter# 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.

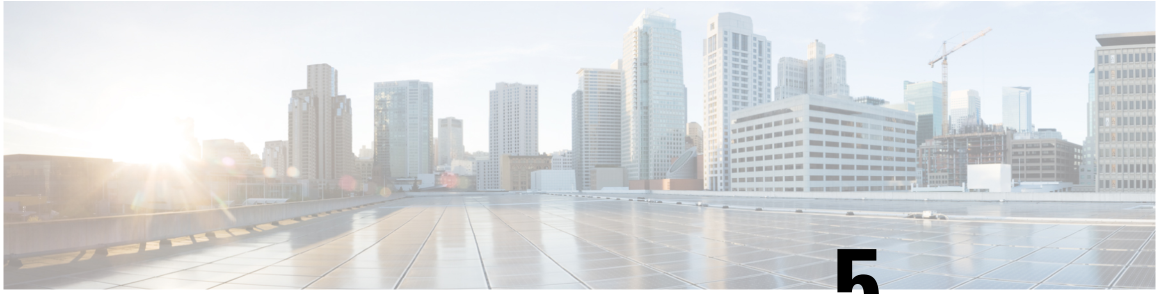
```
Router #sh isis segment-routing prefix-sid-map receive
IS-IS Level-1 receive prefix-sid maps:
Host                 Prefix               SID Index    Range        Flags
utah                 16.16.16.16/32       101          1
                     16.16.16.17/32       102          1            Attached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

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, seeUsing 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).

Using OSPF with Segment Routing

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

- 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 st 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 directed 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.

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 OSF 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: 0xFEF8
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 0xC
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*]

Additional References

Related Documents

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

Standards and RFCs

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

MIBs

MIB	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: http://www.cisco.com/go/mibs

Technical Assistance

Description	Link
The Cisco Support website provides extensive online resources, including documentation and tools for troubleshooting and resolving technical issues with Cisco products and technologies. 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. Access to most tools on the Cisco Support website requires a Cisco.com user ID and password.	http://www.cisco.com/cisco/web/support/index.html

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.

- SR-TE LSP Instantiation, on page 54
- SR-TE and TI-LFA, on page 62

Restrictions for SR-TE

- SR-TE statistics counters are not supported on the Cisco ASR 900 RSP2 module.
- The Cisco ASR 900 routers with RSP2 module support five label stacks.
- 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 46.

Note All five labels cannot be part of the SR-TE Tunnel label stack. One label must be a service label.

- Cisco ASR 900 routers with RSP3 module support seven-label stack with recirculation.
- 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.
- 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 statistics counter:
  - Statistics counter for ECMP to first hop node is not supported.
  - Statistics counters for traffic steered over PFP policy is not supported.
  - Statistics counters for Labelled Traffic over PDP SR policy is not supported.
  - For L2VPN prefixes going over more than one auto-route tunnels, an ECMP LB path is formed with SR-TE. Since this is a preselection, only one path is picked up for L2VPN.
  - Per-SR Policy, per TC Aggregate traffic counter (POL.TC)—Per-traffic class aggregate counters per-SR policy is not supported.
  - Per-SR Policy labelled steered aggregate traffic counter (POL.BSID)—Per-binding SID aggregate counters per SR-policy is not supported.
  - Per-SR Policy, Per-Segment-List aggregate traffic counter (POL.SL)—Multiple segment list is not supported.

- Per-SR Policy, Per-Segment-List, Per-Path aggregate traffic counter (POL.SL.INT.E)—Multiple segment lists and tunnels pointing to ECMP next-hop is not supported.
- Statistics counters under BGP-PIC-EDGE is not supported in this release.

Segment Routing Recirculation for TI-LFA

Prior to Cisco IOS XE Gibraltar 16.11.1, the Cisco ASR 900 routers with RSP2 module support five label stacks, whereas the Cisco ASR 900 routers with RSP3 module support four label stacks from which only one label can be used as the SR-TE transport label and one label as TI-LFA label

However, effective Cisco IOS XE Gibraltar 16.11.1, the Cisco ASR 900 routers with RSP3 module support a sven-label stack and more than one SR-TE label in both, the primary and backup paths.

This increase in 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.
- VPLS over SR-TE is not supported.

- ECMP over SR-TE is not supported.
- 6PE and 6VPE are not supported with three and four transport labels.
- 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

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 Tunnel1
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#
```

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:

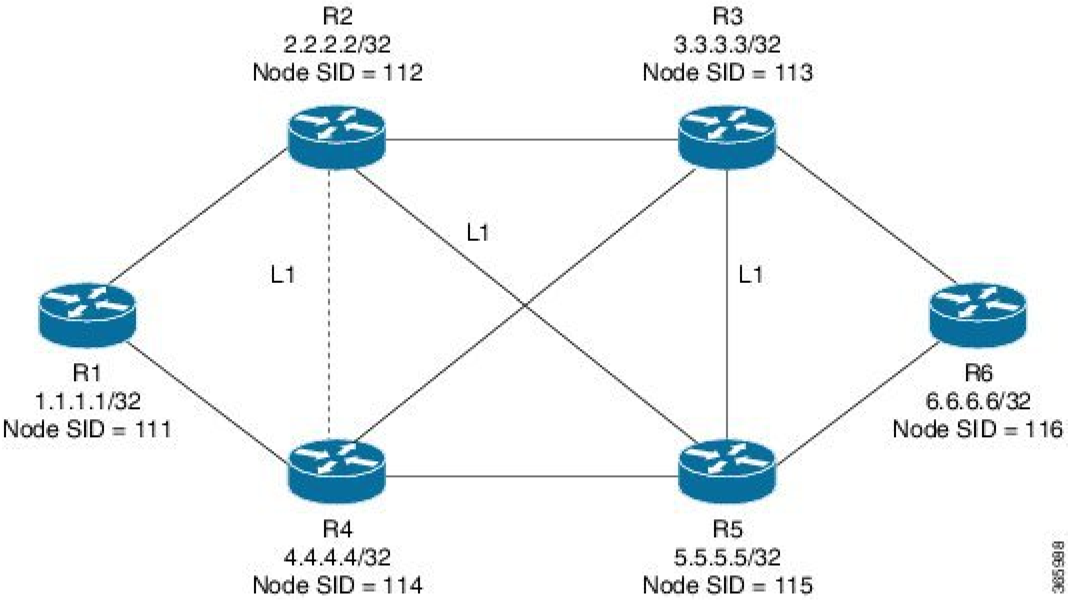
- IP addresses

- 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 4is 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                           (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: 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                           (Tunnel1) 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                           (Tunnel1) 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                           (Tunnel1) 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                           (Tunnel1) 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 IGPs 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 Tunnel1
 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                           (Tunnel1) 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 Tunnel1
 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: R1_t1                           (Tunnel1) 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

ECMP across multiple SR-TE tunnels is not supported.

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 tunnel1
Name: R1_t1                           (Tunnel1) 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 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 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 IGPs *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 IGPs update 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, hwcn]
    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 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

- 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

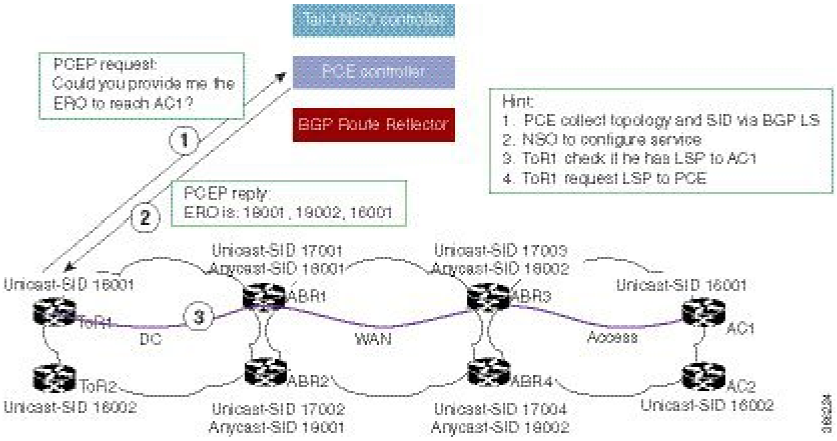
The SR TE On demand LSP feature provides the ability to connect Metro access rings via a static route to the destination.

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
```

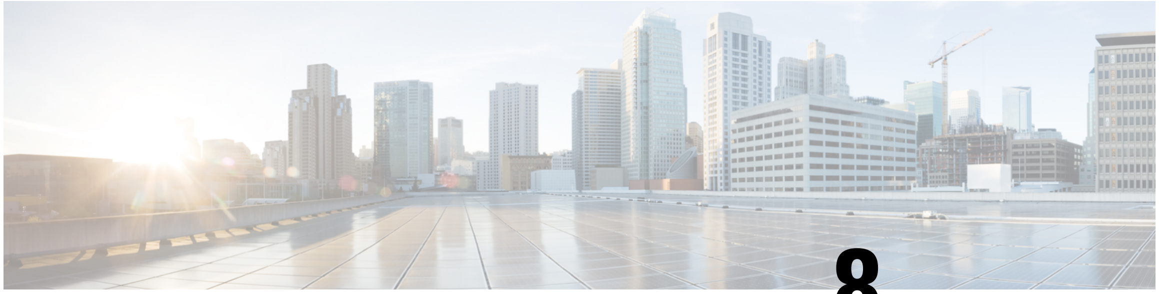
```
PATH path1 (strict source route, path complete, generation 5)
    1: next-label 16055
    2: next-label 16070
    3: next-label 16010
PATH path2 (strict source route, path complete, generation 9)
    1: next-label 16070
    2: next-label 16010
Device# show running-config | i ip route
ip route  10.10.10.10 255.255.255.255 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 8

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.

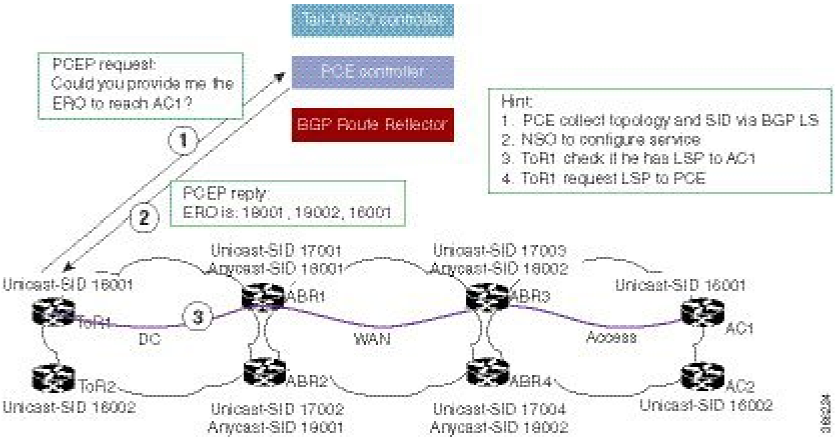
Restrictions for SR-TE On-Demand Next Hop

- ODN anycast SID is not supported.
- ODN for IPv6 is not supported.
- Only BGP-LS is supported on the Cisco RSP3 Module. BGP-LS with ODN is supported on the Cisco RSP2 Module.
- 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 SPFs 200 msecs
 Maximum wait time between two consecutive SPFs 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 msecs
Minimum hold time between two consecutive SPFs 200 msecs
Maximum wait time between two consecutive SPFs 5000 msecs
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
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 thewards 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

- On the Cisco ASR 903 routers with RSP3 module, the total stack limit is four labels without recirculation. However, with recirculation, total stack limit is seven labels.

```
Router# show mpls outlabel-stack reservation
 Application   Outlabels        Status
               Reserved
   bgp            0            COMPLETE
   transport      0            COMPLETE
   vpn            1            COMPLETE    ---□ 1 Label is used for VPN service
 Platform Max OutLabels:4
 Platform Max Primary Recirc OutLabels:0
 Platform Max Repair Recirc OutLabels:3
 Total Available Primary Path OutLabels including recirc Labels:3
 Total Available Repair Path OutLabels including recirc Labels:6
 Reg apps: OSPF-1(3) MPLS-TE(3)

Router#
```

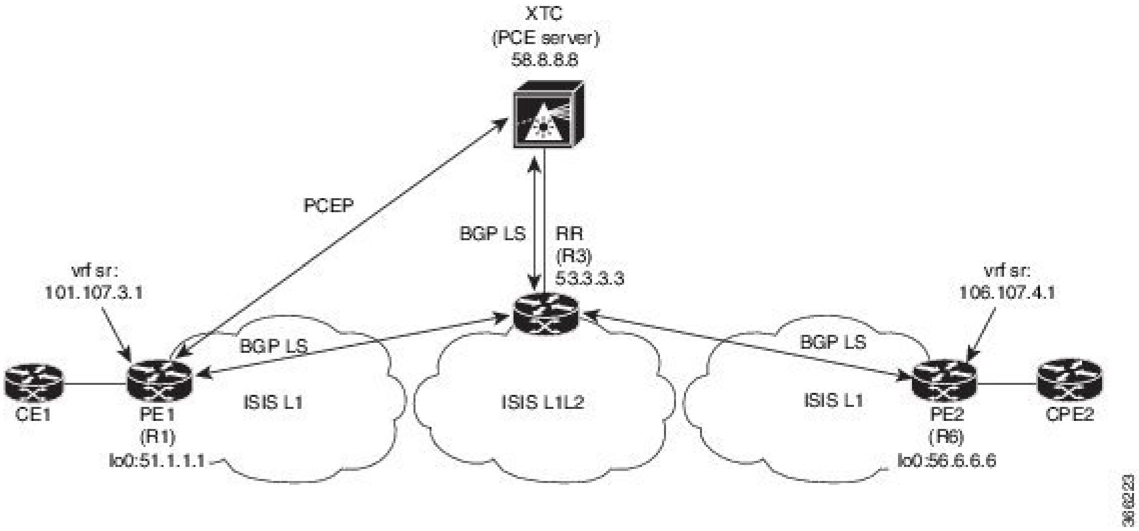
- The maximum number of SR-TE ODN policies supported is 500.
- Static SR-TE policy is not supported.
- SR-TE policy counters are supported on the Cisco ASR 903 routers with RSP3 module only if the number of transport labels used for SR-TE policy does not exceed two labels.
- SR-TE policy goes down only 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



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
ispf 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 nolabel/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).

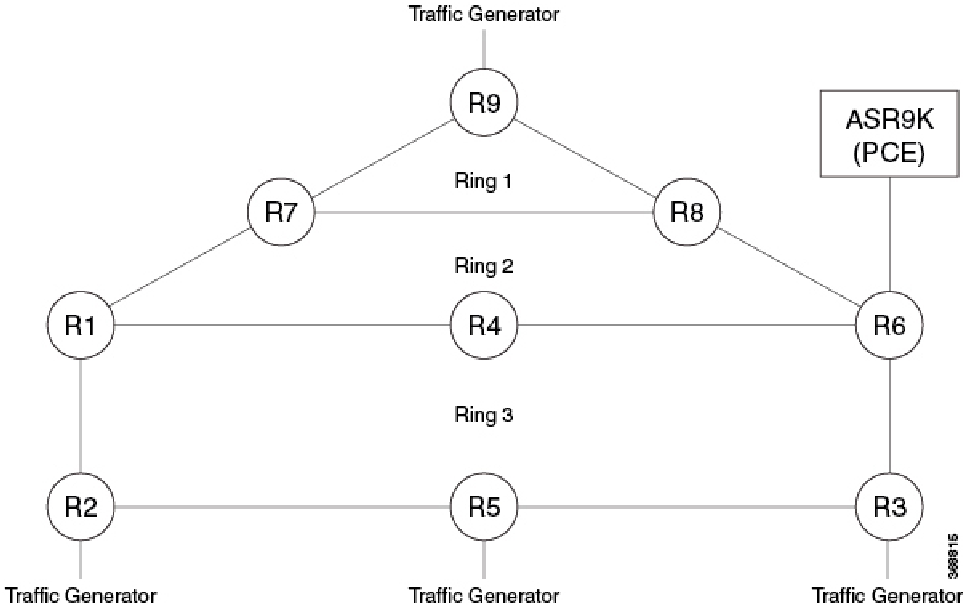
```
Device# show ip route vrf sr
Routing Table: sr
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - 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 Taileend 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
  neighbor 206.206.206.206 send-community both
  neighbor 206.206.206.206 route-map R9_R3_R5_R2_BGP_INTER_DOMAIN out
 exit-address-family
 !
```

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

```
Router(config-srte-odn-path-pref-dyn-metric)# type ?
  igp  Specify IGP metric
  te   Specify TE metric
```

Configuring Affinity Constraint

```
segment-routing traffic-eng
 interface GigabitEthernet0/2/3
  affinity
   name 1
 on-demand color 100
  authorize restrict
   ipv4 prefix-list R9350_BGP_INTER_DOMAIN
  candidate-paths
   preference 1
    constraints
     segments
      dataplane mpls
     !
 affinity  --------------------□ Affinity configuration
      include-any  -----------□ Affinity Type configuration
       name 1  ---------------□ Affinity Name configuration
       !
     !
    dynamic
     pcep
    !
   !
  !
 !
 pcc
  pce address <pce loopback ip> source-address <pcc loopback ip>
 !
 affinity-map   -----------------------□ Affinity Map configuration
   name 1 bit-position 1
```

Configuring Disjointness Constraint

```
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
     !
     affinity
      include-any
       name 1
         !
     !
     association-group   --------------------□ Disjointness configuration
       identifier 1
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
                                 Label 18002, Address 207.207.207.207
                                 Label 21, Address: local 10.10.20.2 remote 10.10.20.1
  Thu Jan 10 00:09:05 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 23:54:42 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 dispointness 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 9

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

- 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>]]
[{dsmap | ddmap [l2ecmp]} [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>}]
```

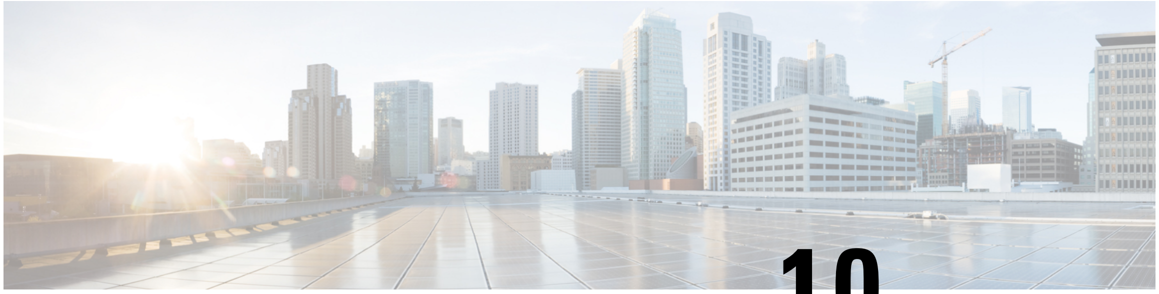
For more information, see traceroute mpls .

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
```

```
Device# show mpls forwarding-table
Local     Outgoing  Prefix          Bytes Label  Outgoing  Next Hop
Label     Label     or Tunnel Id    Switched     interface
16        Pop Label 3333.3333.0000-Et1/0-30.1.1.3  \
        0            Et1/0     30.1.1.3
17        Pop Label 5555.5555.5555-Et1/1-90.1.1.5  \
        0            Et1/1     90.1.1.5
18        Pop Label 3333.3333.0253-Et0/2-102.102.102.2  \
        0             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 10

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

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 Tunnel1
 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: d1_t1                        (Tunnel1) 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: d1_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: d1_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: d1_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: d1_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 11

Segment Routing uLoop Avoidance

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

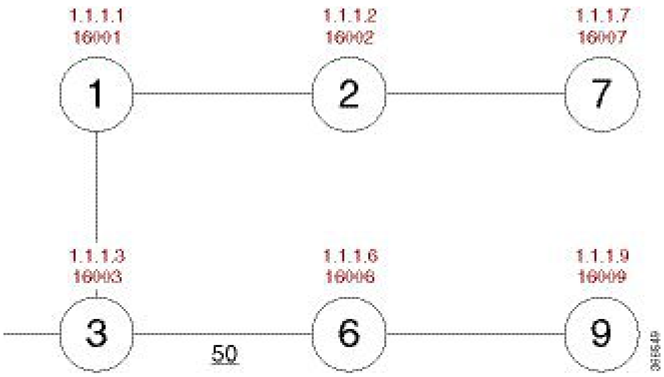
When changes occur in a network topology because of the failure or restoration of a link or a network device, IP Fast Reroute enables rapid network convergence by moving traffic to precomputed backup paths until regular convergence mechanisms move traffic to a newly computed best path, also known as 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



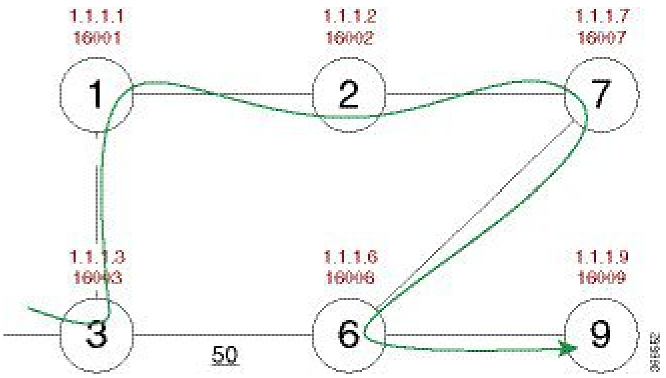
The assumptions in this example are as follows:

- 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.

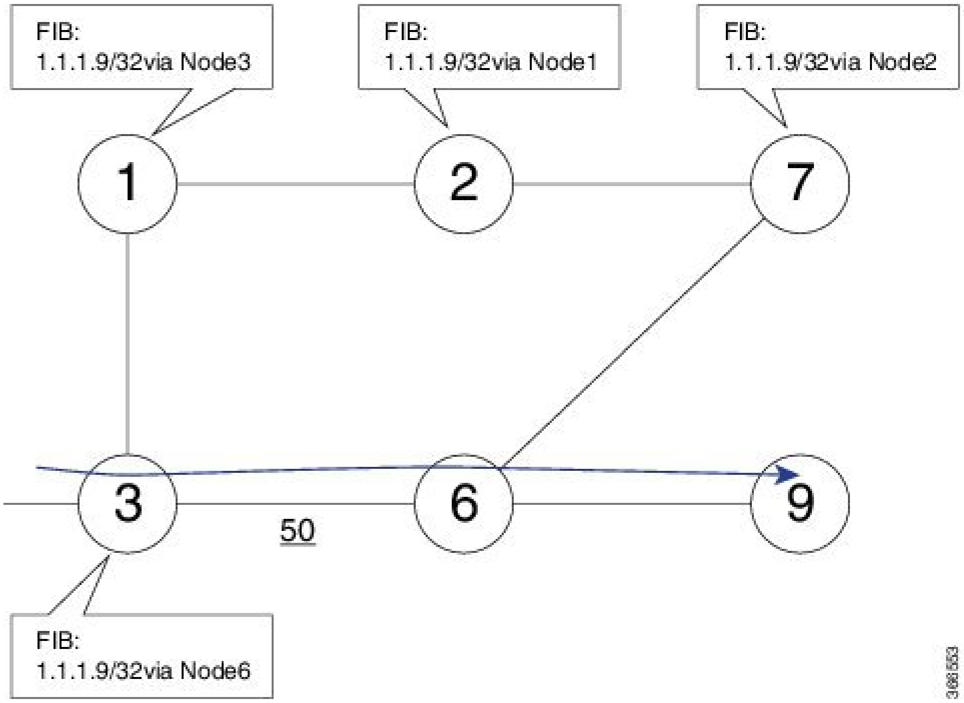
If a link is established between Node 6 and Node 7, the shortest path for a packet from Node 3 to Node 9 would be Node 1, Node 2, Node 7, and Node 6 before the packet reaches the destination, Node 9.

Figure 11: Microloop Example Topology—Shortest Path



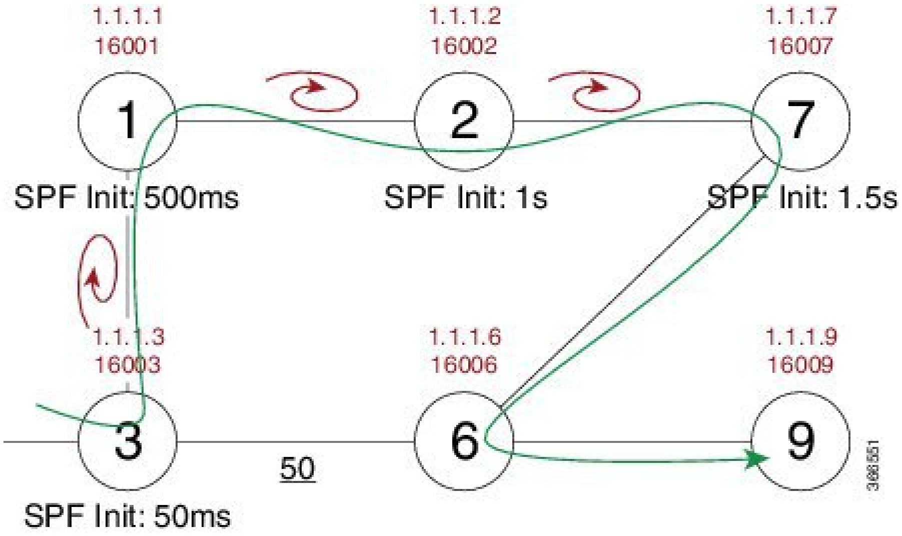
The following figure shows the Forwarding Information Base (FIB) table in each node before the link between Node 6 and Node 7 is established. The FIB entry contains the prefix of the destination node (Node 9) and the next hop.

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

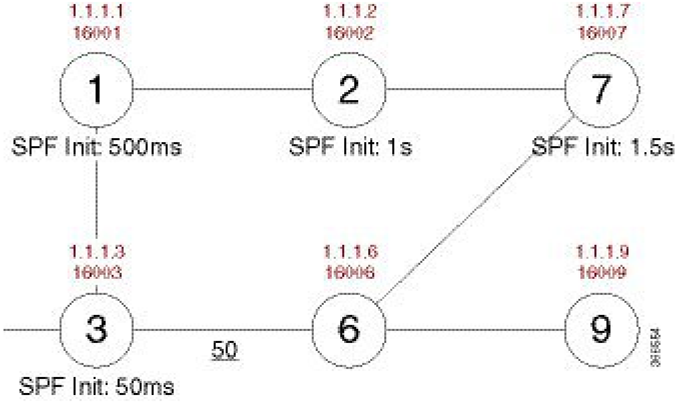


Adding the SPF convergence delay, microloop results in a loss of connectivity for 1.5 seconds, which is the convergence duration specified for node 7.

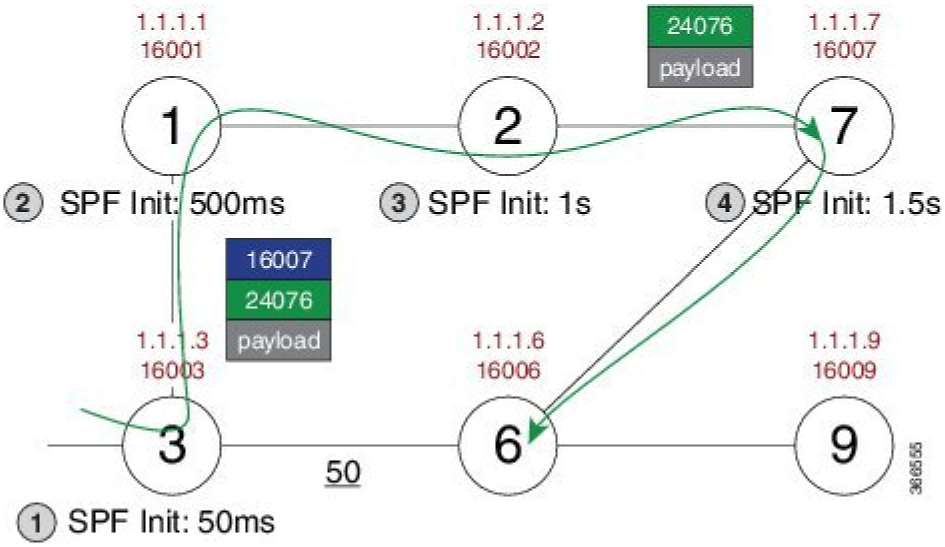
How Segment Routing Prevents Microloops?

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_reqed[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_reqed[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
```