Segment Routing Configuration Guide, Cisco IOS XE Everest 16.5

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

CONTENTS

CHAPTER 15

CHAPTER 16

CHAPTER 1

Read Me First

Important Information about Cisco IOS XE 16

Effective Cisco IOS XE Release 3.7.0E for Catalyst Switching and Cisco IOS XE Release 3.17S (for Access and Edge Routing) the two releases evolve (merge) into a single version of converged release—the Cisco IOS XE 16—providing one release covering the extensive range of access and edge products in the Switching and Routing portfolio.

Feature Information

Use Cisco Feature Navigator to find information about feature support, platform support, and Cisco software image support. An account on Cisco.com is not required.

Related References

- Cisco IOS Command References, All Releases

Obtaining Documentation and Submitting a Service Request

- To receive timely, relevant information from Cisco, sign up at Cisco Profile Manager.
- To get the business impact you’re looking for with the technologies that matter, visit Cisco Services.
- To submit a service request, visit Cisco Support.
- To discover and browse secure, validated enterprise-class apps, products, solutions and services, visit Cisco Marketplace.
- To obtain general networking, training, and certification titles, visit Cisco Press.
- To find warranty information for a specific product or product family, access Cisco Warranty Finder.

CHAPTER 2

Introduction to Segment Routing

This chapter introduces the concept of Segment Routing (SR).

Overview of Segment Routing

Segment Routing (SR) is a flexible, scalable way of doing source routing. The source chooses a path and encodes it in the packet header as an ordered list of segments. Segments are identifier for any type of instruction. Each segment is identified by the segment ID (SID) consisting of a flat unsigned 32-bit integer. Segment instruction can be:

- Go to node N using the shortest path
- Go to node N over the shortest path to node M and then follow links Layer 1, Layer 2, and Layer 3
- Apply service S

With segment routing, the network no longer needs to maintain a per-application and per-flow state. Instead, it obeys the forwarding instructions provided in the packet.

Segment Routing relies on a small number of extensions to Cisco Intermediate System-to-Intermediate System (IS-IS) and Open Shortest Path First (OSPF) protocols. It can operate with an MPLS (Multiprotocol Label Switching) or an IPv6 data plane, and it integrates with the rich multi service capabilities of MPLS, including Layer 3 VPN (L3VPN), Virtual Private Wire Service (VPWS), Virtual Private LAN Service (VPLS), and Ethernet VPN (EVPN).

Segment routing can be directly applied to the Multiprotocol Label Switching (MPLS) architecture with no change in the forwarding plane. Segment routing utilizes the network bandwidth more effectively than traditional MPLS networks and offers lower latency. A segment is encoded as an MPLS label. An ordered list of segments is encoded as a stack of labels. The segment to process is on the top of the stack. The related label is popped from the stack, after the completion of a segment.

Segment routing can be applied to the IPv6 architecture with a new type of routing extension header. A segment is encoded as an IPv6 address. An ordered list of segments is encoded as an ordered list of IPv6 addresses in the routing extension header. The segment to process is indicated by a pointer in the routing extension header. The pointer is incremented, after the completion of a segment.

Segment Routing provides automatic traffic protection without any topological restrictions. The network protects traffic against link and node failures without requiring additional signaling in the network. Existing IP fast re-route (FRR) technology, in combination with the explicit routing capabilities in Segment Routing guarantees full protection coverage with optimum backup paths. Traffic protection does not impose any additional signaling requirements.

How Segment Routing Works

A router in a Segment Routing network is capable of selecting any path to forward traffic, whether it is explicit or 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, an adjacency segment ID is locally significant and represents 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:

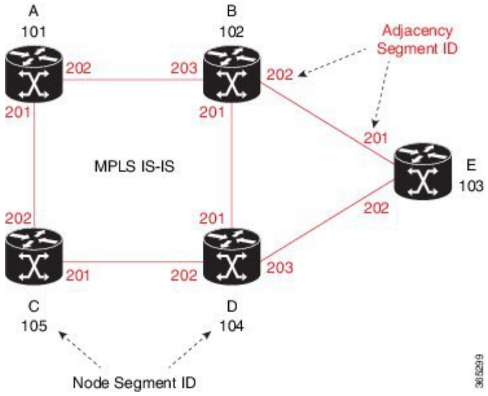
- **Prefix SID**— A segment ID that contains an IP address prefix calculated by an IGP in the service provider core network. Prefix SIDs are globally unique. A prefix segment represents the shortest path (as computed by IGP) to reach a specific prefix; a node segment is a special prefix segment that is bound to the loopback address of a node. 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.

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

Examples for Segment Routing

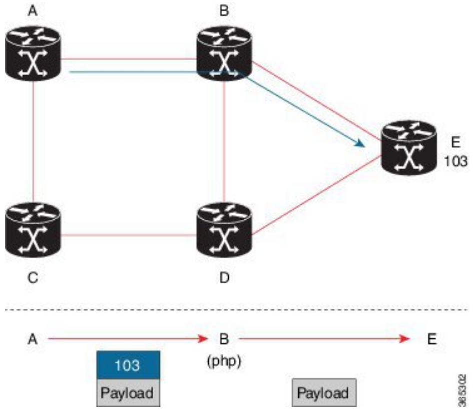
The following figure illustrates an MPLS network with five routers using Segment Routing, IS-IS, a label range of 100 to 199 for node IDs, and 200 and higher for adjacency IDs. IS-IS would distribute IP prefix reachability alongside segment ID (the MPLS label) across the network.

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



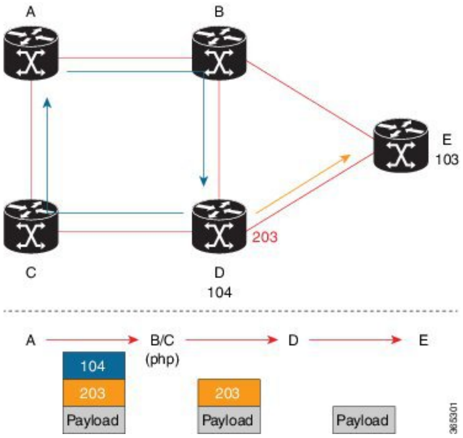
In the previous example, any router sending traffic to router E would push label 103 (router E node segment identifier) to forward traffic using the IS-IS shortest path. The MPLS label-swapping operation at each hop preserves label 103 until the packet arrives at E (Figure 2). On the other hand, adjacency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

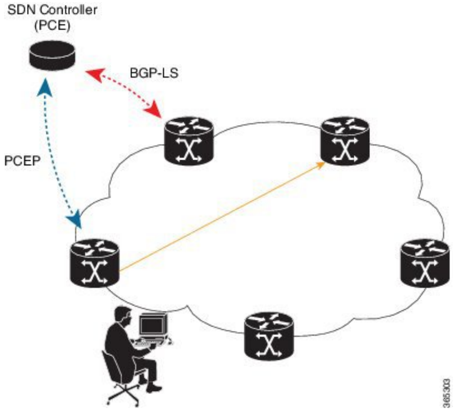
Figure 3: Router E Destination Path



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 very 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. Below 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 (ISIS 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. In case of link or node failures 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 LSP's 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 via plane 1" and "Flow 2 injected in node A towards node Z must go via 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 tail end 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 interoperable with existing MPLS control and data planes and can be implemented in an existing deployment.

Segment Routing Global Block

Segment Routing Global Block (SRGB) is the range of labels reserved for segment routing. SRGB is local property of an segment routing node. In MPLS, architecture, SRGB is the set of local labels reserved for global segments. In segment routing, each node can be configured with a different SRGB value and hence the absolute SID value associated to an IGP Prefix Segment can change from node to node.

The SRGB default value is 16000 to 23999. The SRGB can be configured as follows:

```
Device(config)# router isis 1
Device(config-isis)#segment-routing global-block 45000 55000
```

The SRGB label value is calculated as follows:

- If the platform supports 1000000 labels or more, the SRGB value is from 900000 to 900000 + 2^16.
- If the platform supports less than 1000000 labels, the SRGB value is the last 2^16 labels.

Restrictions:

- The SRGB size cannot be more than 2^16.
- The SRGB upper bound 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.

Segment Routing Global Block

This chapter explains the concept of creating a block of labels reserved for a router using segment routing. This block of reserved labels is known as the Segment Routing Global Block (SRGB).

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. No specific configuration is required to enable adj-SIDs. Once segment routing is enabled over IS-IS for an address-family, for any interface that IS-IS runs over, the address-family automatically allocates an adj-SID towards every neighbor out of that interface.

Note Only IPV4 address-family supports allocating adj-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.

A prefix SID is allocated from the Segment Routing Global Block (SRGB). The prefix SID value translates to a local MPLS label, whose value is calculated as below:

- If the platform supports 1000000 labels or more, then the MPLS label corresponding to the prefix SID value is 900000 + *sid-value*.
- If the platform supports less than 1000000 labels, then the MPLS label corresponding to the prefix SID value is *maximum-supported-label-value* - 2^16 + *sid-value*.

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, in the platform supporting 1000000 MPLS labels or more 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 9000010 16010 to the prefix 1.0.0.1/32.

BGP Prefix Segment Identifiers

Segments associated with a BGP prefix are known as BGP Prefix-SIDs.

- BGP Prefix-SIDs are always global within a Segment Routing or BGP domain
- BGP Prefix-SIDs identifies an instruction to forward the packet over ECMP-aware best path computed by BGP for a given prefix

Segment Routing requires BGP speaker to be configured with a Segment Routing Global block (SRGB). Generally, SRGB is configured as a range of labels, SRGB = [SR_S, SR_E].

- SR_S = Start of the range
- SR_E = End of the range

Each prefix is assigned with its own unique label index.

In the following example, a BGP route policy, set label index, is defined using the route-policy **name** command.

Configure the Segment Routing Global Block (SRGB) in BGP. If the route label path has a label-index attribute and SRGB is configured, then local label route is allocated from SRGB. If label-index is added to redistributed routes using route-policy, then BGP presents label-index as an attribute with the route.

```
router bgp 100
 bgp log-neighbor-changes
 neighbor 192.0.2.1 remote-as 100
 neighbor 192.0.2.1 update-source Loopback0
 neighbor 192.0.23.3 remote-as 300
 !
 address-family ipv4
  segment-routing mpls
  neighbor 192.0.2.1 activate
  neighbor 192.0.2.1 send-label
  neighbor 192.0.23.3 activate
 exit-address-family
```

Additional References for Segment Routing

Related Documents

Related Topic	Document Title
Videos	• Introduction to Cisco Segment Routing (YouTube) • Introduction to Cisco Segment Routing (CCO)

Feature Information for Introduction to Segment Routing

The following table provides release information about the feature or features described in this module. This table lists only the software release that introduced support for a given feature in a given software release train. Unless noted otherwise, subsequent releases of that software release train also support that feature.

Use Cisco Feature Navigator to find information about platform support and Cisco software image support. To access Cisco Feature Navigator, go to www.cisco.com/go/cfn. An account on Cisco.com is not required.

Table 1: Feature Information for Introduction to Segment Routing

Feature Name	Releases	Feature Information
Introduction to Segment Routing	Cisco IOS XE Release 3.16S Cisco IOS XE Fuji 16.7.1	Segment Routing (SR) is a flexible, scalable way of doing source routing. In Cisco IOS XE Fuji 16.7.1, this feature is supported on Cisco 4000 Series Integrated Service Routers.

CHAPTER 3

Segment Routing With IS-IS v4 Node SID

This chapter describes how Segment Routing (SR) works with IS-IS.

Restrictions for Segment Routing With IS-IS v4 Node SID

- 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).
- Effective Cisco IOS-XE Release 3.16, ISIS supports segment routing for IPv4 only.

Information About Segment Routing IS-IS v4 Node SID

Segment Routing IS-IS v4 Node SID

Segment Routing relies on a small number of extensions to Cisco Intermediate System-to-Intermediate System (IS-IS) and Open Shortest Path First (OSPF) protocols. 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. 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.

Prefix-SID Received in Label Switched Path 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) in the same way as BGP downloads per prefix VPN labels, only if 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.

Note

- For SIDs that do not fit in the specified SID range, labels are not used when updating the RIB. For the cases, where SID fits 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 all the next hops do not support segment routing, ISIS treats the instance similar to mismatched labels assigned to the same prefix. That means, 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 3.17, 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 3.18, 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 3.18, 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 3.17, 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.

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 3.16 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

The command **sr-label-preferred** allows the forwarding interface to prefer the segment routing labels over LDP labels for all prefixes in a topology.

Segment Routing -Traffic Engineering 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.

How to Configure Segment Routing —IS-IS v4 Node SID

Configuring Segment Routing

Before you begin

Before configuring IS-IS to support segment routing you must first configure the segment routing feature in global configuration mode.

SUMMARY STEPS

1. **enable**
2. **configure terminal**

3. segment-routing mpls
4. connected-prefix-sid-map
5. address-family ipv4
6. 1.1.1.1/32 index 100 range 1
7. exit-address-family

DETAILED STEPS

	Command or Action	Purpose
Step 1	**enable** **Example:** `Device# enable`	Enables privileged EXEC mode. • Enter your password if prompted.
Step 2	**configure terminal** **Example:** `Device# configure terminal`	Enters global configuration mode.
Step 3	segment-routing mpls **Example:** `Device(config-sr)# segment-routing mpls`	Enables the segment feature using the mpls data plane.
Step 4	connected-prefix-sid-map **Example:** `Device(config-srmpls)# connected-prefix-sid-map`	Enters a sub-mode where you can configure address-family specific mappings for local prefixes and SIDs.
Step 5	address-family ipv4 **Example:** `Device(config-srmpls-conn)# address-family ipv4`	Specifies IPv4 address prefixes.
Step 6	1.1.1.1/32 index 100 range 1 **Example:** `Device(config-srmpls-conn-af)# 1.1.1.1/32 100 range 1`	Associates SID 100 with the address 1.1.1.1/32.
Step 7	exit-address-family **Example:** `Device(config-srmpls-conn-af)# exit-address-family`	Exits the address family.

Configuring Segment Routing on IS-IS Network

Before you begin

Before you configure segment routing on IS-IS network, IS-IS must be enabled on your network.

SUMMARY STEPS

1. router isis
2. net network-entity-title
3. metric-style wide
4. **segment-routing** mpls
5. exit
6. show isis segment-routing

DETAILED STEPS

	Command or Action	Purpose
Step 1	router isis **Example:** `Device(config-router)# router isis`	Enables the IS-IS routing protocol and enters router configuration mode.
Step 2	net network-entity-title **Example:** `Device(config-router)# net 49.0000.0000.0003.00`	Configures network entity titles (NETs) for the routing instance.
Step 3	metric-style wide **Example:** `Device(config-router)# metric-style wide`	Configures the device to generate and accept only wide link metrics.
Step 4	**segment-routing** mpls **Example:** `Device(config-router)# segment-routing mpls`	Configures segment routing operation state.
Step 5	exit **Example:** `Device(config-router)# exit`	Exits segment routing mode and returns to the configuration terminal mode.
Step 6	show isis segment-routing **Example:** `Device# show is-is segment-routing`	Displays the current state of the IS-IS segment routing.

Example

The following example displays output from the show isis segment-routing state command for the segment routing under IS-IS:

```
Device# show isis segment-routing

ISIS protocol is registered with MFI
ISIS MFI Client ID:0x63
Tag 1 - Segment-Routing:
   SR State:SR_ENABLED
   Number of SRGB:1
   SRGB Start:16000, Range:8000, srgb_handle:0x4500AED0, srgb_state: created
   Address-family IPv4 unicast SR is configured
     Operational state:Enabled
```

Configuring Prefix-SID for IS-IS

This task explains how to configure prefix segment identifier (SID) index under each interface.

Before you begin

Segment routing must be enabled on the corresponding address family.

SUMMARY STEPS

1. enable
2. configure terminal
3. segment-routing mpls
4. connected-prefix-sid-map
5. address-family ipv4
6. 1.1.1.1/32 index 100 range 1
7. exit

DETAILED STEPS

	Command or Action	Purpose
Step 1	enable **Example:** `Device# enable`	Enables privileged EXEC mode.
Step 2	configure terminal **Example:** `Device# configure terminal`	Enters global configuration mode.
Step 3	segment-routing mpls **Example:**	Configures segment routing mpls mode.

	Command or Action	Purpose
	`Device(config)# segment-routing mpls`	
Step 4	connected-prefix-sid-map **Example:** `Device(config-srmpls)# connected-prefix-sid-map`	Enters a sub-mode where you can configure address-family specific mappings for local prefixes and SIDs.
Step 5	address-family ipv4 **Example:** `Device(config-srmpls-conn)# address-family ipv4`	Specifies the IPv4 address family and enters router address family configuration mode.
Step 6	1.1.1.1/32 index 100 range 1 **Example:** `Device(config-srmpls-conn-af)# 1.1.1.1/32 100 range 1`	Associates SID 100 with the address 1.1.1.1/32.
Step 7	exit **Example:** `Device(config-router)# exit`	Exits segment routing mode and returns to the configuration terminal mode.

Configuring Prefix Attribute N-flag-clear

By default, a flag called N-flag is set by IS-IS when advertising a SID which is associated with a loopback address. If you wish to clear this flag add explicit configuration.

SUMMARY STEPS

1. **enable**
2. **configure terminal**
3. interface loopback3
4. isis prefix n-flag-clear

DETAILED STEPS

	Command or Action	Purpose
Step 1	**enable** **Example:** `Device# enable`	Enables privileged EXEC mode. • Enter your password if prompted.
Step 2	**configure terminal** **Example:**	Enters global configuration mode.

	Command or Action	Purpose
	`Device# configure terminal`	
Step 3	interface loopback3 **Example:** `Device(config)# interface loopback3`	Specifies the interface loopback.
Step 4	isis prefix n-flag-clear **Example:** `Device(config-if)# isis prefix n-flag-clear`	Clears the prefix N-flag.

Configuring Explicit Null Attribute

To disable penultimate-hop-popping (PHP) and add explicit-Null label, explicit-null option needs to be specified. Once the option is given, IS-IS sets the E flag in the prefix-SID sub TLV.

By default, a flag called E-flag (Explicit-Null flag) is set to 0 by ISIS when advertising a Prefix SID which is associated with a loopback address. If you wish to set this flag add explicit configuration.

SUMMARY STEPS

1. **enable**
2. **configure terminal**
3. segment-routing mpls
4. set-attributes
5. address-family ipv4
6. explicit-null
7. exit-address-family

DETAILED STEPS

	Command or Action	Purpose
Step 1	**enable** **Example:** `Device# enable`	Enables privileged EXEC mode. • Enter your password if prompted.
Step 2	**configure terminal** **Example:** `Device# configure terminal`	Enters global configuration mode.
Step 3	segment-routing mpls **Example:**	Configures segment routing mpls mode.

	Command or Action	Purpose
	`Device(config)# segment-routing mpls`	
Step 4	set-attributes **Example:** `Device(config-srmpls)# set-attributes`	Sets the attribute.
Step 5	address-family ipv4 **Example:** `Device(config-srmpls-attr)# address-family ipv4`	Specifies the IPv4 address family and enters router address family configuration mode.
Step 6	explicit-null **Example:** `Device(config-srmpls-attr-af)# explicit-null`	Specifies the explicit-null.
Step 7	exit-address-family **Example:** `Device(config-srmpls-attr-af)# exit-address-family`	Exits the address family.

Configuring Segment Routing Label Distribution Protocol Preference

SUMMARY STEPS

1. **enable**
2. **configure terminal**
3. segment-routing mpls
4. set-attributes
5. address-family ipv4
6. sr-label-preferred
7. exit-address-family

DETAILED STEPS

	Command or Action	Purpose
Step 1	**enable** **Example:** `Device# enable`	Enables privileged EXEC mode. • Enter your password if prompted.
Step 2	**configure terminal** **Example:**	Enters global configuration mode.

	Command or Action	Purpose
	`Device# configure terminal`	
Step 3	segment-routing mpls **Example:** `Device(config)# segment-routing mpls`	Configures segment routing mpls mode.
Step 4	set-attributes **Example:** `Device(config-srmpls)# set-attributes`	Sets the attribute.
Step 5	address-family ipv4 **Example:** `Device(config-srmpls-attr)# address-family ipv4`	Specifies the IPv4 address family and enters router address family configuration mode.
Step 6	sr-label-preferred **Example:** `Device(config-srmpls-attr-af)# sr-label-preferred`	Specifies SR label to be preferred over the LDP.
Step 7	exit-address-family **Example:** `Device(config-srmpls-attr-af)# exit-address-family`	Exits the address family.

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 local and remote mapping entries.

The following command disables the prefix-sid-mapping client functionality and it is configured on the receiver side.

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

Configuring IS-IS SID Binding TLV Domain Flooding

By default, the IS-IS SRMS server does not flood SID binding entries within the routing domain. From Cisco IOS-XE Release 3.18, the optional keyword **domain-wide** is added 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.

Configuration Examples for Segment Routing —IS-IS v4 Node SID

Example: Configuring Segment Routing on IS-IS Network

The following example shows how to configure prefix segment identifier (SID) index under each interface:

```
Device(config)#segment-routing mpls
 Device(config-srmpls)#connected-prefix-sid-map
  Device(config-srmpls-conn)#address-family ipv4
   Device(config-srmpls-conn-af)#10.1.2.2/32 index 2 range 1
  Device(config-srmpls-conn-af)#exit-address-family
 Device(config-srmpls-conn-af)#end
```

Example: Configuring Explicit Null Attribute

The following is an example for configuring explicit null attribute:

```
Device(config)# segment-routing mpls
Device(config-srmpls)# set-attributes
 Device(config-srmpls-attr)# address-family ipv4
  Device(config-srmpls-attr-af)# explicit-null
 Device (config-srmpls-attr-af)# exit-address-family
```

Additional References for Segment Routing With IS-IS v4 Node SID

Related Documents

Related Topic	Document Title
Cisco IOS commands	Cisco IOS Master Command List, All Releases http://www.cisco.com/c/en/us/td/docs/ios-xml/ios/mcl/allreleasemcl/all-book.html
IP Routing ISIS commands	Cisco IOS IP Routing ISIS commands http://www.cisco.com/c/en/us/td/docs/ios-xml/ios/mcl/allreleasemcl/all-book.html

RFCs

RFC	Title
RFC4971	Intermediate System to Intermediate System (IS-IS) Extensions for Advertising Router Information
RFC5305	IS-IS Extensions for Traffic Engineering. Defines the advertisement of router IDs for IPv4.
RFC6119	IPv6 Traffic Engineering in IS-IS. Defines the advertisement of router IDs for IPv6.

Feature Information for Segment Routing—IS-IS v4 Node SID

The following table provides release information about the feature or features described in this module. This table lists only the software release that introduced support for a given feature in a given software release train. Unless noted otherwise, subsequent releases of that software release train also support that feature.

Use Cisco Feature Navigator to find information about platform support and Cisco software image support. To access Cisco Feature Navigator, go to www.cisco.com/go/cfn. An account on Cisco.com is not required.

Table 2: Feature Information for Segment Routing—IS-IS v4 Node SID

Feature Name	Releases	Feature Information
Segment Routing—IS-IS v4 Node SID	Cisco IOS XE Release 3.16S Cisco IOS XE Fuji 16.7.1	The Segment Routing—ISIS v4 node SID feature provides support for segment routing on IS-IS networks. The following commands were introduced or modified: **connected-prefix-sid-map**, **show isis segment-routing**, **isis prefix n-flag-clear**, **explicit-null** In Cisco IOS XE Fuji 16.7.1, this feature is supported on Cisco 4000 Series Integrated Service Routers.

CHAPTER 4

IS-IS Link-protection Topology Independent Loop Free Alternate Fast Reroute

This document describes the functionalities and IS-IS implementation of IP Fast Re-Route feature (IPFRR) using Segment Routing (SR) Topology Independent Loop Free Alternative (TI-LFA) link protection.

Prerequisites for IS-IS Link-protection Topology Independent Loop Free Alternate Fast Reroute

- Enable TI-LFA on all the nodes, before configuring SR-TE for TI-LFA.

```
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
 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
 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
```

- If a microloop gets created between routers in case of primary and secondary path switch over you need to bring down the convergence time. Use the **microloop avoidance rib-update-delay** command to bring down the convergence time:

```
router isis ipfrr
net 49.0001.0120.1201.2012.00
is-type level-2-only
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
```

- Enable MPLS-TE nonstop routing (NSR) and IS-IS nonstop forwarding (NSF) to reduce or minimize traffic loss after a high availability (HA) switch over. Use the **mpls traffic-eng nsr** command in global exec mode.

```
mpls traffic-eng nsr
```

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

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

Information About IS-IS Link-protection Topology Independent Loop Free Alternate Fast Reroute

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 3.18, topology-independent LFA (TI-LFA) is supported on an SR-enabled network.

Topology-Independent Loop Free Alternate

TI-LFA provides supports for 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.

Topology Independent Loop Free Alternate Tie-break

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

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

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

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

Interface Fast Reroute Tiebreakers

Interface fast reroute (FRR) tiebreakers are also needed for TI-LFA node and SRLG protection. When interface and protocol instance FRR tiebreakers both 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.

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 3.18, node-protecting tie-breaker is disabled by default.

How to Configure IS-IS Link-protection Topology Independent Loop Free Alternate Fast Reroute

Perform the following steps to configure Link-protection Topology Independent Loop Free Alternate Fast Reroute.

Configuring Topology Independent Loop Free Alternate Fast Reroute

You can enable TI-LFA using any of the following two methods:

1. **Protocol enablement**—Enables TI-LFA in router isis mode for all IS-IS 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]
```

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

2. **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]
```

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

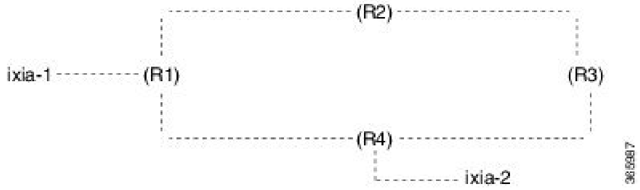
   When both interface and protocol are TI-LFA enabled, the interface configuration takes precedence over the protocol configuration. TI-LFA is disabled by default.

   To disable TI-LFA on a particular interface, use the following command:

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

Configuring Topology Independent Loop Free Alternate With Mapping Server

Consider the following topology to understand the configuration:



- IXIA-2 injects ISIS 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:

```
configure terminal
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

```
configure terminal
!
!
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

```
configure terminal
!
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:

```
configure terminal
!
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
!
```

Examples: Configuring IS-IS Link-protection Topology Independent Loop Free Alternate Fast Reroute

Example 1: In the following example, local LFA is configured with linecard-disjoint and srlg-disjoint tiebreakers. 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 10000
 bfd all-interfaces
```

Example 2—Enable TI-LFA node-protecting tie-breaker on all ISIS level-2 interfaces with priority 100. 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 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 4—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 5—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 6—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 Tiebreaker

To view tiebreakers enabled on the interface, use the following command:

show running all | section interface *interface-name*

To view tiebreakers enabled on the router mode, use the following command:

show running all | section router isis

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 Topology Independent Loop Free Alternate 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
```

Verifying the Segment Routing Traffic Engineering With Topology Independent Loop Free Alternate 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 head-end router:

```
interface Tunnel1
ip unnumbered Loopback1
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
```

Additional References for IS-IS Link-protection Topology Independent Loop Free Alternate Fast Reroute

Related Documents

Related Topic	Document Title
Cisco IOS Commands	Cisco IOS Master Command List, All Releases

Feature Information for IS-IS Link-protection Topology Independent Loop Free Alternate Fast Reroute

Table 3: Feature Information for IS-IS Link-protection Topology Independent Loop Free Alternate Fast Reroute

Feature Name	Releases	Feature Information
IS-IS Link-protection Topology Independent Loop Free Alternate Fast Reroute	Cisco IOS XE Everest 16.4.1 Cisco IOS XE Fuji 16.7.1	The following commands were introduced or modified: **fast-reroute ti-lfa {level-1 | level-2} [maximum-metric value]** , isis fast-reroute ti-lfa protection level-1 disable, **isis fast-reroute ti-lfa protection {level-1 | level-2} [maximum-metric value]**, **show running all | section interface interface-name**, **show running all | section router isis**. In Cisco IOS XE Fuji 16.7.1, this feature is supported on Cisco 4000 Series Integrated Service Routers.

CHAPTER 5

Segment Routing Traffic Engineering With IS-IS

This chapter describes how segment routing traffic engineering (SR-TE) can be implemented using IS-IS.

Restrictions for Segment Routing-Traffic Engineering With IS-IS

- SR-TE is not supported on broadcast interfaces; it is supported only point-to-point interfaces.
- The Cisco ASR routers support only a specific number of labels imposed on an outgoing packet. If the number of labels are greater than the specified number, the SR-TE tunnel creation fails. The Cisco ASR1000 routers support a maximum of 16 labels.
- Only one instance of protocol should be enabled for TE at a given point of time.
- You can use the verbatim keyword only on a label-switched path (LSP) that is configured with the explicit path option.
- Re-optimization is unsupported on the verbatim LSP.

Information About Segment Routing Traffic Engineering With IS-IS

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 head-end 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

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 head-end 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 tail-end without a transport label. However, in some cases, it is desirable that the packet arrive at the tail-end with explicit-null label, and in such case, the head-end will impose an explicit-null label at the top of the label stack.

SR-TE LSP Path Verification

SR-TE tunnel functionality requires that the head-end perform initial verification of the tunnel path as well as the subsequent tracking of the reachability of the tunnel tail-end 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 head-end validates the path of an SR-TE LSP for connectivity against the TE topology. MPLS-TE head-end checks if links corresponding to the adjacency SIDs are connected in the TE topology.

For newly-instantiated SR-TE LSPs, if the head-end 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 head-end detects a discontinuity on any link, the head-end assumes a fault has occurred on that link. In this case, the local repair protection, such as the IP FRR, come 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 head-end 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 head-end 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 head-end 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 head-end. 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 head-end has full visibility over the LSP path, and validates the path to the ultimate LSP destination. However, for inter-area LSPs, the head-end has partial visibility over the LSP path—only up to the first ABR. In this case, the head-end 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 head-end, 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 head-end 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 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 head-end 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 head-end validates that the specified SR path is compliant with the configured affinity. This necessitates 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.

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 head-end 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 commands below. 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
!
```

Verbatim Path Support

MPLS TE LSPs usually require that all the nodes in the network are TE aware which means that they have IGP extensions to TE in place. However, some network administrators want the ability to build TE LSPs to traverse nodes that do not support IGP extensions to TE, but that do support RSVP extensions to TE. Verbatim LSPs are helpful when all or some of the intermediate nodes in a network do not support IGP extensions for TE.

When this feature is enabled, the IP explicit path is not checked against the TE topology database. Since the TE topology database is not verified, a Path message with IP explicit path information is routed using the shortest path first (SPF) algorithm for IP routing.

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 Single Tunnel

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 the head-end or any midpoint traversed node along the SR-TE LSP path.

Load Balancing on Multiple Tunnels

Multiple TE tunnels can be used as next-hop paths for routes to specific IP prefixes either by configuring static route on multiple tunnels, or auto-route announcing multiple parallel tunnels to the same destination. In such cases, the tunnels share the traffic load equally or load balance traffic on multiple parallel tunnels. It is also possible to allow Unequal Load Balance (UELB) with an explicit per tunnel configuration at the tunnel head-end. In this case, the tunnel load-share is passed from MPLS-TE to forwarding plane.

The tunnel load-share feature continues to work for TE tunnels that instantiate the SR-TE LSPs.

SR-TE Tunnel Re-optimization

TE tunnel re-optimization occurs when the head-end determines that there is a more optimal path available than the one currently used. For example, in case of a failure along the SR-TE LSP path, the head-end 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 head-end 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 head-end 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 head-end 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
```

SR-TE With Lockdown Option

The **lockdown** option 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
 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-selection metric igp
 tunnel mpls traffic-eng load-share 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:

IP-FRR Local Repair Protection

On an SR-TE LSP head-end 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 head-end 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 head-end 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.

In the event of a failure of the primary SR-TE LSP, 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.

Unnumbered Support

IS-IS description of an unnumbered link does not contain remote interface ID information. The remote interface ID of an unnumbered link is required to include the unnumbered link as part of the SR-TE tunnel.

How to Configure Segment Routing Traffic Engineering With IS-IS

Perform the following steps to configure Segment Routing Traffic Engineering (SR-TE) with IS-IS.

Configuring Path Option for a TE Tunnel

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

```
Device(config)# interface tunnel 100
Device(config-if)# tunnel mpls traffic-eng path-option 1 explicit name foo segment-routing
Device(config-if)# tunnel mpls traffic-eng path-option 2 dynamic segment-routing
Device(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 will be 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

The following SR-TE explicit path hops are supported:

- IP addresses
- MPLS labels
- Mix of IP addresses and MPLS labels

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

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

Note When using IP unnumbered interfaces, you cannot specify next hop address as an explicit path index. It should be node address or label.

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

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

Configuring Affinity on an Interface

Perform the following steps to configure 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
```

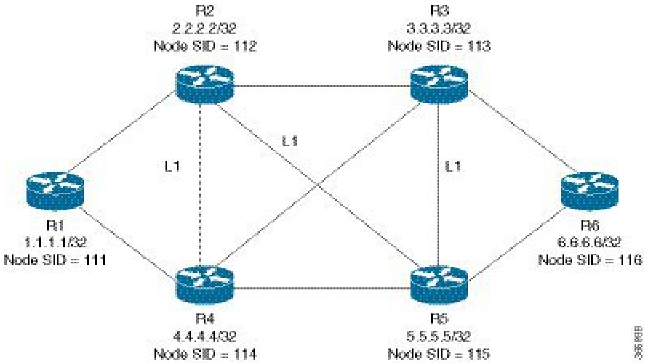
Enabling Verbatim Path Support

To enable verbatim with SR-TE you can use the following example. In the example, tunnel destination 11.11.11.11 is in different area and an explicit path with name multihop is defined with SR-TE path option.

```
R6#
interface Tunnel4
ip unnumbered Loopback66
tunnel mode mpls traffic-eng
tunnel destination 11.11.11.11
tunnel mpls traffic-eng path-option 1 explicit name multihop segment-routing verbatim
!
ip explicit-path name multihop enable
index 1 next-label 16003
index 2 next-label 16002
index 3 next-label 16001
!
End
```

Use Case: Segment Routing Traffic Engineering Basic Configuration

Consider the following topology to understand the SR-TE configuration:



To configure at the head-end router, R1:

```
!
mpls traffic-eng tunnels
!
segment-routing mpls
connected-prefix-sid-map
  address-family ipv4
   1.1.1.1/32 index 111  range 1
  exit-address-family
!
set-attributes
  address-family ipv4
  sr-label-preferred
exit-address-family
!
interface Loopback1
ip address 1.1.1.1 255.255.255.255
ip router isis 1
!
int gig0/0
ip address 11.11.11.1 255.255.255.0
ip router isis 1
mpls traffic-eng tunnels
isis network point-to-point
!
router isis 1
net 49.0001.0010.0100.1001.00
is-type level-1
metric-style wide
segment-routing mpls
segment-routing prefix-sid-map advertise-local
mpls traffic-eng router-id  Loopback1
mpls traffic-eng level-1
!
end
```

To enable SR-TE Explicit path (Node SID based), enable the following CLI on R1:

```
Head end SR-TE configuration R1#
!
interface tunnel1
ip unnumbered Loopback1
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 Node_PATH segment-routing
!
ip explicit-path name Node_PATH
 next-label  16114
next-label  16115
next-label  16116
```

To verify proper operation of SR-TE tunnel 1 on R1 enable the following CLI:

```
Tunnel verification on (R1)# show mpls traffic-eng tun tun 1 detail
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 Node_PATH (Basis for Setup)
  Config Parameters:
    Bandwidth: 0        kbps (Global)  Priority: 6  6   Affinity: 0x0/0xFFFF
    Metric Type: IGP (interface)
Verbatim: disabled
Number of LSP IDs (Tun_Instances) used: 1815
    Current LSP: [ID: 1815]
      Uptime: 2 seconds
Removal Trigger: configuration changed
    Segment-Routing Path Info (isis  level-1)
    Segment0[Node]: 4.4.4.4, Label: 16114
    Segment1[Node]: 5.5.5.5, Label: 16115
    Segment2[Node]: 6.6.6.6, Label: 16116
```

To configure at the tail-end router, R6:

```
interface GigabitEthernet2
ip address 100.101.1.1 255.255.255.0
ip router isis 1
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
 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 load-share 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
 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 load-share 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
 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 load-share 10
```

Note In the case of mixed path, IP next-hop cannot be used after using Node SIDs in the path. The following path will not be valid:

```
ip explicit-path name MIXED_PATH enable
next-label 115
next-label 116
next-address 2.2.2.2
```

Dynamic Path SR-TE Tunnel 4

Consider that tunnel 4is based on adjacency SIDs

```
interface Tunnel4
 ip unnumbered Loopback1
 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 load-share 10
end
```

Dynamic Path SR-TE Tunnel 5

Consider that tunnel 5 is based on Node SIDs

```
 interface Tunnel5
 ip unnumbered Loopback1
 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 load-share 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
```

Verifying Verbatim Path Support

To verify proper operation and SR-TE tunnel state use following CLI:

```
R6#sh mpl traffic-eng tunnels tunnel 4

Name: R6_t4                               (Tunnel4) Destination: 11.11.11.11
  Status:
    Admin: up         Oper: up     Path: valid       Signalling: connected
    path option 1, (SEGMENT-ROUTING) type explicit (verbatim) multihop (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: 16 minutes, 40 seconds
    Time since path change: 13 minutes, 6 seconds
    Number of LSP IDs (Tun_Instances) used: 13
  Current LSP: [ID: 13]
    Uptime: 13 minutes, 6 seconds
    Selection: reoptimization
  Prior LSP: [ID: 12]
    ID: path option unknown
    Removal Trigger: configuration changed (severe)
Tun_Instance: 13
Segment-Routing Path Info (IGP information is not used)
  Segment0[First Hop]: 0.0.0.0, Label: 16003
  Segment1[ - ]: Label: 16002
  Segment2[ - ]: Label: 16001
```

Additional References for Segment Routing Traffic Engineering With IS-IS

Related Documents

Related Topic	Document Title
Cisco IOS Commands	Cisco IOS Master Command List, All Releases

Feature Information for Segment Routing -Traffic Engineering With IS-IS

The following table provides release information about the feature or features described in this module. This table lists only the software release that introduced support for a given feature in a given software release train. Unless noted otherwise, subsequent releases of that software release train also support that feature.

Use Cisco Feature Navigator to find information about platform support and Cisco software image support. To access Cisco Feature Navigator, go to www.cisco.com/go/cfn. An account on Cisco.com is not required.

Table 4: Feature Information for Segment Routing -Traffic Engineering With IS-IS

Feature Name	Releases	Feature Information
Segment Routing-Traffic Engineering With IS-IS	Cisco IOS XE Everest 16.4.1 Cisco IOS XE Fuji 16.7.1	The following commands were introduced or modified: **mpls traffic-eng nsr**, **show mpls traffic-eng tunnels tunnel1**, **show isis fast-reroute ti-lfa tunnel**, **show frr-manager client client-name ISIS interfaces detail**, **show ip cef 6.6.6.6 internal** In Cisco IOS XE Fuji 16.7.1, this feature is supported on Cisco 4000 Series Integrated Service Routers.

CHAPTER 6

Segment Routing With OSPFv2 Node SID

This chapter describes how Segment Routing works with OSPFv2 node SID.

Information About Segment Routing With OSPFv2 Node SID

Segment Routing relies on a small number of extensions to Open Shortest Path First (OSPF) protocols. 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 ospf level enables segment routing for the ospf 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. 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.

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 OSPF 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 OSPF 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 OSPF Router Information Opaque LSA.

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

Prefix-SID Received in Label Switched Path From Remote Routers

OSPF sends the prefix SIDs associated with the connected prefix using the Extended Prefix Sub TLV in its opaque Extended prefix LSA. Prefix SIDs received in a LSA which have got reachability are downloaded to the routing information base (RIB) in the same way as BGP downloads per prefix VPN labels, only if 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.

Note For SIDs that do not fit in the specified SID range, labels are not used when updating the RIB. For the cases, where SID fits in the SID range, but does not fit the next-hop neighbor SID range, remote label associated with that path is not installed.

Segment Routing Adjacency SID Advertisement

Effective with Cisco IOS-XE Release 3.17, OSPF 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 Extended Opaque Link LSA.

OSPF allocates the adjacency SID for each OSPF neighbor if the OSPF adjacency which are in two way or in FULL state. OSPF allocates the adjacency SID only if the Segment Routing is enabled. The label for adjacency SID is dynamically allocated by the system. This eliminates the chances of misconfiguration, as this has got only the local significance.

Multiple Adjacency-SIDs

Effective with Cisco IOS-XE Release 16.3, multiple adjacency-SIDs are supported. For each OSPF adjacency, OSPF allots to Adj SIDs, unprotected and protected Adj-SIDs which are carried in the extended link LSAs. The protected adjacency SID (or back up Adj-SID) is allocated and advertised only when FRR is enabled on the router and also on the interface where SR is enabled on the system. When FRR or SR is disabled, the protected Adj-SID is released.

The persistence of protected adj-SID in forwarding plane is supported. When the primary link is down, OSPF delays the release of its backup Adj-SID until the delay timer (30 sec) expires. This allows the forwarding plane to continue to forward the traffic through the backup path until the router is converged.

The allocated and advertised backup Adj-SIDs can be displayed in the output of **show ip ospf neighbor detail** and **show ip ospf segment-routing protected-adjacencies command**.

Segment Routing Mapping Server

Segment Routing Mapping Server (SRMS) allows configuration and maintenance of the Prefix-SID mapping policy entries. Effective with Cisco IOS-XE Release 3.17, 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

When a router installs 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.

SRGB Range Changes

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

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

MPLS Forwarding on an Interface

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

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

Conflict Handling of SID Entries

When there is a conflict between the SID entries and the associated prefix entries use any of the following methods to resolve the conflict:

- When the system receives two SID entries for the same prefix, then the prefix received by higher router ID is treated as the SID corresponding to the prefix. The prefix is installed with the SID entry which was advertised by the higher router ID.
- When the system receives two SID entries one by OSPF protocol and the other by IS-IS protocol, then the SID entry received by OSPF protocol is treated as valid SID. The prefix is installed with the SID entry which was received by OSPF protocol.
- When two prefixes are advertised with the same SID entry, then the prefix which is advertised by the higher router ID is installed with the SID entry and the other prefix is installed without any SID entry.

In an ideal situation, each prefix should have unique SID entries assigned.

How to Configure Segment Routing With OSPFv2 Node SID

Perform the following steps to configure segment routing with OSPFv2 node SID.

Configuring Segment Routing With OSPF

Before you begin

Before configuring OSPF to support segment routing you must first configure the segment routing feature in global configuration mode.

SUMMARY STEPS

1. **enable**
2. **configure terminal**
3. segment-routing mpls
4. connected-prefix-sid-map
5. address-family ipv4
6. 1.1.1.1/32 index 100 range 1
7. exit-address-family

DETAILED STEPS

	Command or Action	Purpose
Step 1	**enable** **Example:** `Device# enable`	Enables privileged EXEC mode. • Enter your password if prompted.
Step 2	**configure terminal** **Example:** `Device# configure terminal`	Enters global configuration mode.
Step 3	segment-routing mpls **Example:** `Device(config-sr)# segment-routing mpls`	Enables the segment feature using the mpls data plane.
Step 4	connected-prefix-sid-map **Example:** `Device(config-srmpls)# connected-prefix-sid-map`	Enters a sub-mode where you can configure address-family specific mappings for local prefixes and SIDs.
Step 5	address-family ipv4 **Example:**	Specifies IPv4 address prefixes.

	Command or Action	Purpose
	`Device(config-srmpls-conn)# address-family ipv4`	
Step 6	1.1.1.1/32 index 100 range 1 **Example:** `Device(config-srmpls-conn-af)# 1.1.1.1/32 100 range 1`	Associates SID 100 with the address 1.1.1.1/32.
Step 7	exit-address-family **Example:** `Device(config-srmpls-conn-af)# exit-address-family`	Exits the address family.

Configuring Segment Routing on OSPF Network

Before you begin

Before you configure segment routing on OSPF network, OSPF must be enabled on your network.

SUMMARY STEPS

1. **router ospf 10**
2. **router-id**<*id*>
3. **segment-routing mpls**
4. **segment-routing area** <*area id*> **mpls**
5. **show ip ospf 10 segment-routing**

DETAILED STEPS

	Command or Action	Purpose
Step 1	**router ospf 10** **Example:** `Device(config)# router ospf 10`	Enables the OSPF mode.
Step 2	**router-id**<*id*> **Example:** `Device(config-router)# router-id 1.0.0.0`	Configures OSPF routes.
Step 3	**segment-routing mpls** **Example:** `Device(config-router)# segment-routing mpls`	Configures segment routing mpls mode.

	Command or Action	Purpose
Step 4	**segment-routing area** *<area id>* **mpls** **Example:** `Device(config-router) # segment-routing area 0 mpls`	Configures segment routing mpls mode in a specific area.

Step 5	**show ip ospf 10 segment-routing**
	Example:
	`Device# show ip ospf 10 segment-routing`
		Shows the output for configuring SR under OSPF.
		The following example displays output from the show ip ospf segment-routing state command for the segment routing under OSPF:
		`Device#show ip ospf 10 segment-routing`
		`           OSPF Router with ID (0.0.0.1) (Process ID 10)`
		`Global segment-routing state: Enabled`
		`Segment Routing enabled:`
		`           Area        Topology name   Forwarding`
		`              0                 Base      MPLS`
		`              1                 Base      MPLS`
		`SR Attributes`
		`    Prefer non-SR (LDP) Labels`
		`    Do not advertise Explicit Null`
		`Local MPLS label block (SRGB):`
		`    Range: 16000 - 23999`
		`    State: Created`
		`Registered with SR App, client handle: 3`
		`  Connected map notifications active (handle 0x4), bitmask 0x1`
		`  Active policy map notifications active (handle 0x5), bitmask 0xC`
		`Registered with MPLS, client-id: 100`
		`Bind Retry timer not running`
		`Adj Label Bind Retry timer not running`

Configuring Prefix-SID for OSPF

This task explains how to configure prefix segment identifier (SID) index under each interface.

Before you begin

Segment routing must be enabled on the corresponding address family.

SUMMARY STEPS

1. enable
2. configure terminal
3. segment-routing mpls

4. connected-prefix-sid-map
5. address-family ipv4
6. 1.1.1.1/32 index 100 range 1
7. exit

DETAILED STEPS

	Command or Action	Purpose
Step 1	enable **Example:** `Device# enable`	Enables privileged EXEC mode.
Step 2	configure terminal **Example:** `Device# configure terminal`	Enters global configuration mode.
Step 3	segment-routing mpls **Example:** `Device(config)# segment-routing mpls`	Configures segment routing mpls mode.
Step 4	connected-prefix-sid-map **Example:** `Device(config-srmpls)# connected-prefix-sid-map`	Enters a sub-mode where you can configure address-family specific mappings for local prefixes and SIDs.
Step 5	address-family ipv4 **Example:** `Device(config-srmpls-conn)# address-family ipv4`	Specifies the IPv4 address family and enters router address family configuration mode.
Step 6	1.1.1.1/32 index 100 range 1 **Example:** `Device(config-srmpls-conn-af)# 1.1.1.1/32 100 range 1`	Associates SID 100 with the address 1.1.1.1/32.
Step 7	exit **Example:** `Device(config-router)# exit`	Exits segment routing mode and returns to the configuration terminal mode.

Configuring Prefix Attribute N-flag-clear

OSPF advertises prefix SIDs via Extended Prefix TLV in its opaque LSAs. It carries flags for the prefix and one of them is N flag (Node) indicating that any traffic sent along to the prefix is destined to the router originating the LSA. This flag typically marks host routes of router's loopback.

SUMMARY STEPS

1. **enable**
2. **configure terminal**
3. interface loopback3
4. ip ospf prefix-attributes n-flag-clear

DETAILED STEPS

	Command or Action	Purpose
Step 1	**enable** **Example:** `Device# enable`	Enables privileged EXEC mode. • Enter your password if prompted.
Step 2	**configure terminal** **Example:** `Device# configure terminal`	Enters global configuration mode.
Step 3	interface loopback3 **Example:** `Device(config)# interface loopback3`	Specifies the interface loopback.
Step 4	ip ospf prefix-attributes n-flag-clear **Example:** `Device(config-if)# ip ospf prefix-attributes n-flag-clear`	Clears the prefix N-flag.

Configuring Explicit Null Attribute With OSPF

To disable penultimate-hop-popping (PHP) and add explicit-Null label, explicit-null option needs to be specified. Once the option is given, OSPF sets the E flag in the Extended prefix-SID TLV in its LSAs.

By default, a flag called E-flag (Explicit-Null flag) is set to 0 by OSPF when advertising a Prefix SID which is associated with a loopback address. If you wish to set this flag add explicit configuration.

SUMMARY STEPS

1. **enable**
2. **configure terminal**

3. segment-routing mpls
4. set-attributes
5. address-family ipv4
6. explicit-null
7. exit-address-family

DETAILED STEPS

	Command or Action	Purpose
Step 1	**enable** **Example:** `Device# enable`	Enables privileged EXEC mode. • Enter your password if prompted.
Step 2	**configure terminal** **Example:** `Device# configure terminal`	Enters global configuration mode.
Step 3	segment-routing mpls **Example:** `Device(config)# segment-routing mpls`	Configures segment routing mpls mode.
Step 4	set-attributes **Example:** `Device(config-srmpls)# set-attributes`	Sets the attribute.
Step 5	address-family ipv4 **Example:** `Device(config-srmpls-attr)# address-family ipv4`	Specifies the IPv4 address family and enters router address family configuration mode.
Step 6	explicit-null **Example:** `Device(config-srmpls-attr-af)# explicit-null`	Specifies the explicit-null.
Step 7	exit-address-family **Example:** `Device(config-srmpls-attr-af)# exit-address-family`	Exits the address family.

Configuring Segment Routing Label Distribution Protocol Preference With OSPF

SUMMARY STEPS

1. **enable**
2. **configure terminal**
3. segment-routing mpls
4. set-attributes
5. address-family ipv4
6. sr-label-preferred
7. exit-address-family

DETAILED STEPS

	Command or Action	Purpose
Step 1	**enable** **Example:** `Device# enable`	Enables privileged EXEC mode. • Enter your password if prompted.
Step 2	**configure terminal** **Example:** `Device# configure terminal`	Enters global configuration mode.
Step 3	segment-routing mpls **Example:** `Device(config)# segment-routing mpls`	Configures segment routing mpls mode.
Step 4	set-attributes **Example:** `Device(config-srmpls)# set-attributes`	Sets the attribute.
Step 5	address-family ipv4 **Example:** `Device(config-srmpls-attr)# address-family ipv4`	Specifies the IPv4 address family and enters router address family configuration mode.
Step 6	sr-label-preferred **Example:** `Device(config-srmpls-attr-af)# sr-label-preferred`	Specifies SR label to be preferred over the LDP.

	Command or Action	Purpose
Step 7	exit-address-family **Example:** `Device(config-srmpls-attr-af)# exit-address-family`	Exits the address family.

Configuring OSPF SRMS

The following command enables the OSPF SRMS and allows OSPF to advertise local mapping entries. OSPF does not send remote entries to the SRMS library. However, OSPF 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 OSPF SRMS Client

By default, the OSPF SRMS client mode is enabled. OSPF always sends remote prefix-sid-mapping entries received through LSAs, 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 and it is configured on the receiver side.

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

Additional References for Segment Routing With OSPFv2 Node SID

Related Documents

Related Topic	Document Title
Cisco IOS commands	Cisco IOS Master Command List, All Releases http://www.cisco.com/c/en/us/td/docs/ios-xml/ios/mcl/allreleasemcl/all-book.html
IP Routing ISIS commands	Cisco IOS IP Routing ISIS commands http://www.cisco.com/c/en/us/td/docs/ios-xml/ios/mcl/allreleasemcl/all-book.html

Feature Information for Segment Routing With OSPFv2 Node SID

Table 5: Feature Information for Segment Routing With OSPFv2 Node SID

Feature Name	Releases	Feature Information
Segment Routing With OSPF	Cisco IOS XE Release 3.16S Cisco IOS XE Fuji 16.7.1	The Segment Routing OSPFv2 node SID feature provides support for segment routing on OSPF networks. The following commands were introduced or modified: **connected-prefix-sid-map**, **show ip ospf 10 segment-routing**, **sr-label-preferred**, **ip ospf prefix-attributes n-flag-clear** . In Cisco IOS XE Fuji 16.7.1, this feature is supported on Cisco 4000 Series Integrated Service Routers.

CHAPTER 7

OSPFv2 Link-protection Topology Independent Loop Free Alternate Fast Reroute

This document describes OSPFv2 implementation of IP Fast Re-Route Feature (IP FRR) using TI -LFA (Topology Independent Loop Free Alternative).

Restrictions for Topology Independent Loop Free Alternate Fast Reroute

- TI-LFA is supported only on OSPFv2.
- TI-LFA tunnels are created only if the router supports SR and it is configured with prefix SID. The prefix (or) node SID can be configured as connected SID (or) advertised using the SRMS (Segment Routing Mapping Server).
- TI-LFA is not supported on OSPF point to multi point interfaces.
- TI-LFA does not support Multi Topology Routing (MTR).
- TI-LFA does not create the repair path using virtual link, sham link (or) TE tunnels.
- TI-LFA tunnel is constructed and programmed by explicitly specifying the node (or) set of repair nodes through which the tunnel needs to traverse.

Information About OSPFv2 Link-Protection Topology Independent Loop Free Alternate Fast Reroute

Topology-Independent Loop-Free Alternate (TI-LFA) uses segment routing to provide link, node, and Shared Risk Link Groups (SRLG) protection in topologies where other fast reroute techniques, such as RLFA (Remote Loop Free Alternative) cannot provide protection. The goal of TI-LFA is to reduce the packet loss that results while routers converge after a topology change due to a link failure. Rapid failure repair (< 50 msec) is achieved through the use of pre-calculated backup paths that are loop-free and safe to use until the distributed network convergence process is completed.

The following are the major benefits of using TI-LFA:

- Provides 100% coverage for all the prefixes and within 50-msec link and node protection.
- Prevents transient congestion and sub-optimal routing by leveraging on the post-convergence path.
- Protects Label Distribution Protocol (LDP) and IP traffic as well.

IP Fast Reroute and Remote Loop Free Alternate

IP Fast Reroute (FRR) is a set of techniques that allow rerouting the IP traffic around a failed link or failed node in the network within a very short time (<50ms). One of the techniques that is used is Loop Free Alternates (LFA), which is implemented using OSPF protocol. OSPF currently supports per-prefix directly connected LFA and remote LFA (RLFA). The problem with these LFA algorithms is the topology dependency; the LFA algorithms cannot find a loop-free alternate path through the network for all the topologies.

The per-prefix directly connected LFA (also known as DLFA) provides loop-free alternate path for most triangular topologies, but does not provide good coverage for rectangular or circular topologies. The Remote LFA implementation (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 P-Space (set of nodes reachable from calculating node without crossing the protected link) and Q-Space (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 node for the protected traffic. RLFA forms targeted LDP session to the PQ node and form the RLFA tunnel. But for the topologies where P and Q-Spaces are disjoint, R-LFA does not provide protection for those prefixes.

Topology Independent Fast Reroute

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

TI-LFA algorithm allows to create a full explicit path through the network. Using fully specified path may lead to issues in larger topologies due to the number of segments along the path. Specifying the whole path is however not necessary, 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. The TI-LFA algorithm constructs a SR tunnel as the repair path. TI-LFA tunnel is constructed and programmed by explicitly specifying the node (or) set of repair nodes through which the tunnel needs to traverse. The traffic is carried on the tunnel (when the primary path fails) which is also on the post convergence path.

Topology-Independent Loop Free Alternate

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, topology-independent LFA (TI-LFA) is supported on an SR-enabled network and provides the following support:

- **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 is not 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**—TI-LFA provides complete coverage of all prefixes, up to the platform’s maximum supported labels.
- **P2P and Broadcast interfaces for the protected link**—TI-LFA protects P2P and broadcast interfaces.
- **Asymmetrical links**—The OSPF metrics between the neighbors are not the same.
- **Multi-homed (anycast) prefix protection**—The same prefix may be originated by multiple nodes and TI-LFA protects the anycast prefixes also by providing post convergence repair path.
- **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.

Topology Independent Loop Free Alternate Tie-break

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.
- **Node-protecting**—Prefers node protecting repair path.
- **SRLG-disjoint**—Prefers SRLG disjoint repair path.
- **Load-sharing**—Distributes repair paths equally among links and prefixes.

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.

- **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.

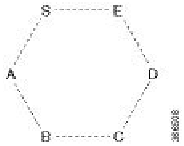
Effective with Cisco IOS-XE Release 3.18, node-protecting tie-breaker is disabled by default. Tie-breaker default and explicit tie-breaker on the same interface are mutually exclusive. The following tie-breakers are enabled by default on all LFAs:

- linecard-disjoint
- lowest-backup-metric
- SRLG-disjoint

P-Space

The set of routers that can be reached from S on the shortest path tree without traversing S-E is termed the P-space of S with respect to the link S-E.

Figure 5: A Simple Ring Topology



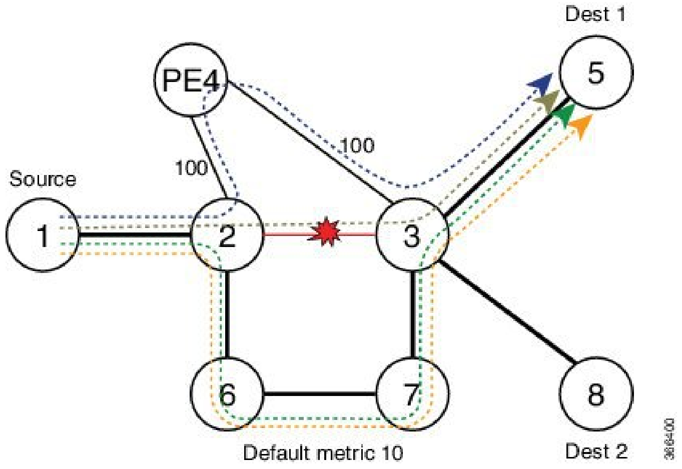
Q-Space

The set of routers from which the node E can be reached, by normal forwarding without traversing the link S-E, is termed the Q-space of E with respect to the link S-E.

Post-Convergence Path

Post convergence path is the path that OSPF uses after the link failure. TI-LFA always calculates the repair path which is the post convergence path. You can plan and dimension the post-convergence path to carry the traffic in the case of failure. TI-LFA enforces the post-convergence path by encoding it as a list of segments. The following figure shows an example of TI-LFA using post convergence path:

Figure 6: TI-LFA Using Post Convergence Path

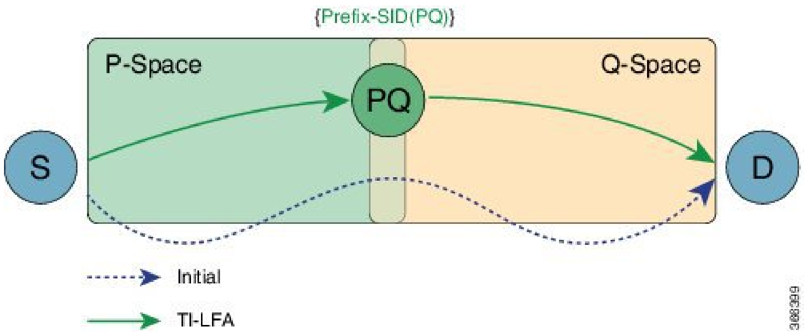


- It protects destination Node 5 on Node 2 against failure of link 2-3.
- Node 2 switches all the traffic destined to Node 5 via core links.

Per-Destination Link Protection

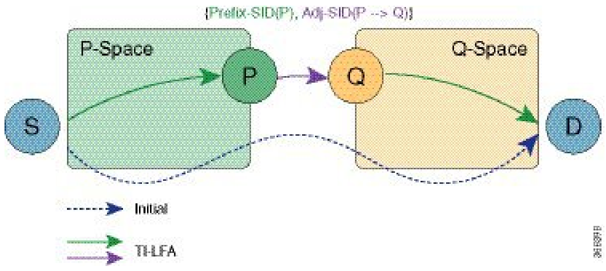
TI-LFA implementation provides per-destination link protection with the number of segments (labels)supported by the underlying hardware. The following figures show the implementation of TI-LFA:

Figure 7: TI-LFA: { Prefix-SID(PQ) }



If PQ is a direct neighbor of S, then no additional segment must be pushed.

Figure 8: TI-LFA: { Prefix-SID(P) , Adj -SID (P -> Q) }



Per Interface Loop Free Alternate Enablement

- TI-LFA can be enabled on an area basis.
- TI-LFA backup path is calculated only if TI-LFA protection is enabled on the primary interface which is to be protected. By default all the interfaces are enabled for protection.
- TI-LFA repair path is restricted by the number of labels supported by the hardware. If hardware supports only 2 labels then TI-LFA repair path can protect only those prefixes which can be protected by 2 or lesser segments. For those prefixes which need more than 2 segment remain unprotected.

Prefix Processing

Once TI-LFA path is calculated for the all the links, prefix processing starts. By default only intra and inter area prefixes are protected. For external prefixes to be protected, you need to enable segment routing globally under the OSPF level.

The primary and repair path should be of the same route type for the prefixes that are protected, that means, if the intra area needs to be protected then the TI-LFA repair path also calculates for the same intra area prefix whether the prefix is unique (or) anycast prefix.

Anycast Prefix Processing

OSPF TI-LFA also calculates the repair path for the anycast prefixes. Anycast prefixes (or) dual homed prefixes are the prefixes advertised by more than one routers. They could be intra, inter (or), external prefixes. The calculation of TI-LFA repair path for anycast prefixes is as below:

- Assume the prefix P1 is advertised by the routers R1 and R2. The prefix advertised by both the routers should be of the same route type, that is, both R1 and R2 should advertise the prefix as intra area prefix (or inter or external).
- Take the primary path is calculated towards R1 due to the lesser cost.
- When TI-LFA calculates the back up path, it calculates the post convergence path. So, post convergence path need not be towards R1. If the cost to reach R2 (in the post convergence) is shorter, then TI-LFA algorithm chooses the post convergence path towards R2. TI-LFA tunnel is formed towards R2.
- When R2 un-advertises the prefix, then the TI-LFA algorithm is re-calculated towards R1 for the repair path.

Per-Prefix Loop Free Alternate Tie-Break

IP FRR has the following tie break rules in the order given below. If you have more than one repair path available to choose the best path from, the following tie-break rules are applied. If more than one path matches all the tie break rules, then all the paths are used as repair paths.

- **Post Convergence**: Prefers backup path which is the post convergence path. This is enabled by default and user can not modify this.
- **Primary-path**: Prefers backup path from ECMP set.
- **Interface-disjoint**: Point-to-point interfaces have no alternate next hop for rerouting if the primary gateway fails. You can set the interface-disjoint attribute to prevent selection of such repair paths, thus protecting the interface.

- **Lowest-backup-metric**: Prefers backup path with lowest total metric. This is not applicable for TI-LFA since TI-LFA always chooses the back up path which is lowest cost.
- **LC-disjoint**: Prefers the back up path which is in different line card than that of the primary path.
- **Broadcast-interface-disjoint** : LFA repair paths protect links when a repair path and a protected primary path use different next-hop interfaces. However, on broadcast interfaces if the LFA repair path is computed via the same interface as the primary path and their next-hop gateways are different, in that case the node gets protected, but the link might not be. You can set the broadcast-interface-disjoint attribute to specify that the repair path never crosses the broadcast network the primary path points to, that means, it cannot use the interface and the broadcast network connected to it.
- **Load Sharing**: When more than one repair path matches the above rules, load share the backup paths. This rule also can be modified by the user.

Note The user can alter and define the tiebreak rules according to the requirement. In this way, the user can re-prioritize the sequence and/or remove some of the tie break indexes which are not needed.

Note The Lowest-backup-metric policy is not applicable for TI-LFA since TI-LFA always chooses the lowest back up path only.

You can see the above rules by using the following command:

```
R2#show ip ospf fast-reroute

            OSPF Router with ID (2.2.2.200) (Process ID 10)

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        Low                   No                Yes
    AS external                 Base        Low                   No                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 17:25:51 ago and was running for 3 ms.
```

With the introduction of TI-LFA, the following two tie-break rules are enhanced.

- node-protection
- srlg-protection

The above two tie-break rules are not enabled by default. The user needs to configure the above mentioned tie-break policies.

Node Protection

TI-LFA node protection provides protection from node failures. Node protecting TI-LFA attempts to calculate the post conversion repair path that protects against the failure of a particular next-hop, not just the link to that particular next-hop.

Node protection is used as a tiebreaker in the implementation of the local LFA also. But when it is combined with TI-LFA, the back up path calculated post convergences with node protecting path. Per-Prefix TI-LFA node protection is disabled by default. The IPFRR TI-LFA node protection features is enabled when the corresponding tiebreak is enabled along with TI-LFA feature, that is,

```
router ospf 10
   [no] fast-reroute per-prefix ti-lfa [area <area> [disable]]
   [no] fast-reroute per-prefix tie-break node-protecting index <index>
   [no] fast-reroute per-prefix tie-break node-protecting required index <index>
```

When you enable node protection, all the other tie break rules also need to manually configured. The node protection is built over the link protection.

The difference between **node-protecting** and **node-protecting required** is in selecting the backup path. When you configure **node-protecting required**, then back up which is chosen has to be the path which does not go through the node (which is part of the link which we are protecting). If no such path is available, then no path is chosen as the backup path.

Shared Risk Link Groups Protection

A shared risk link group (SRLG) is a group of next-hop interfaces of repair and protected primary paths that have a high likelihood of failing simultaneously. The OSPFv2 Loop-Free Alternate Fast Reroute feature supports only SRLGs that are locally configured on the computing router. With the introduction of TI LFA, the post convergence path which does not share the SRLG group id with the primary path interface will be chosen. In that way, the user will be sure of the SRLG protection whenever the primary link fails.

The IPFRR TI-LFA SRLG protection features is enabled when the corresponding tiebreak is enabled along with Ti-LFA feature, that is,

```
router ospf 10
   [no] fast-reroute per-prefix ti-lfa [area <area> [disable]]
   [no] fast-reroute per-prefix tie-break srlg index <index>
   [no] fast-reroute per-prefix tie-break srlg required index <index>
```

When you enable SRLG protection, you need to manually configure all the other tie break rules. The difference between **srlg-protecting** and **srlg-protecting required** is in selecting the backup path. When you configure **srlg-protecting required**, then back up which is chosen has to be the path which does not share SRLG ID with the primary link which is protected. If no such path is available, then no path is chosen as the backup path.

Whereas, if you configure **srlg-protecting** alone then if the SRLG protection path is not available, the link protection path is chosen as the backup path. And when the SRLG protection path is available, the switchover happens to the SRLG protection path.

Node-Shared Risk Link Groups Protection

You can configure both node and SRLG protection tie breaks together. This means that the back up path needs to fulfil both the criteria of node protection as well as SRLG protection. 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 shortest path tree (SPT).

To enable node and SRLG protection tie breaks together, use the following command:

```
router ospf 10
   [no] fast-reroute per-prefix ti-lfa [area <area> [disable]]
   [no] fast-reroute per-prefix tie-break node-protecting index <index>
   [no] fast-reroute per-prefix tie-break srlg index <index>
```

The following show command is used to display the tie break policy:

```
R3#show ip ospf fast-reroute

            OSPF Router with ID (3.3.3.33) (Process ID 10)

Loop-free Fast Reroute protected prefixes:

           Area        Topology name   Priority   Remote LFA Enabled     TI-LFA Enabled
              0                 Base        Low                   No                 No
              1                 Base        Low                   No                 No
           1000                 Base        Low                   No                 No
    AS external                 Base        Low                   No                 No

  Repair path selection policy tiebreaks:
      0  post-convergence
     60  node-protecting
     70  srlg
    256  load-sharing

OSPF/RIB notifications:
 Topology Base: Notification Disabled, Callback Not Registered

Last SPF calculation started 00:00:06 ago and was running for 2 ms.
```

How to Configure Topology Independent Loop Free Alternate Fast Reroute

Enabling Topology Independent Loop Free Alternate Fast Reroute

By default, TI-LFA is disabled. You can use protocol enablement to enable TI-LFA.

Protocol enablement: Enables TI-LFA in router OSPF mode for all the OSPF areas. Perform the following steps to enable TI-LFA FRR.

```
[no] fast-reroute per-prefix ti-lfa [ area <area> disable]
```

```
  router ospf <process>
  fast-reroute per-prefix enable area <area> prefix-priority {low | high}
  fast-reroute per-prefix ti-lfa [ area <area> disable]
```

You can also use interface command to enable or disable IP FRR on specific interfaces.

```
interface <interface>
ip ospf fast-reroute per-prefix protection disable
ip ospf fast-reroute per-prefix candidate disable
ip ospf fast-reroute per-prefix protection ti-lfa [disable]
```

Note

- When TI-LFA is configured on the OSPF router and area wide, area specific configuration takes precedence.
- To protect external prefixes, TI-LFA should be enabled globally.

Configuring Topology Independent Loop Free Alternate Fast Reroute

This task describes how to enable per-prefix Topology Independent Loop-Free Alternate (TI-LFA) computation to converge traffic flows around link, node, and SRLG failures. TI-LFA can be configured on instance or area level inherited by lower levels. You can enable or disable per prefix FRR per interface level which is applicable for TI-LFA also.

Before you begin to configure, ensure that the following topology requirements are met:

- Router interfaces are configured as per the topology.
- Routers are configured with OSPF.
- Segment routing is enabled globally as well as under OSPF level.

1. Enables OSPF routing for the specified routing process and enters in router configuration mode.

```
Device(config)# router ospf 10
```

2. Enables FRR.

```
Device(config-router)# fast-reroute per-prefix enable prefix-priority low
```

3. Enables TI-LFA.

```
Device(config-router)# fast-reroute per-prefix ti-lfa
```

4. Enables TI-LFA on the specific area.

```
Device(config-router)# fast-reroute per-prefix ti-lfa area 0
```

5. Exits the TI-LFA mode.

```
Device(config-router)# exit
```

6. Enters the interface mode.

```
Device(config)#interface ethernet 0/0
```

7. If you do not wish to enable FRR on a specific inteface, use the protection disable command.

```
Device(config-if)#ip ospf fast-reroute per-prefix protection disable
```

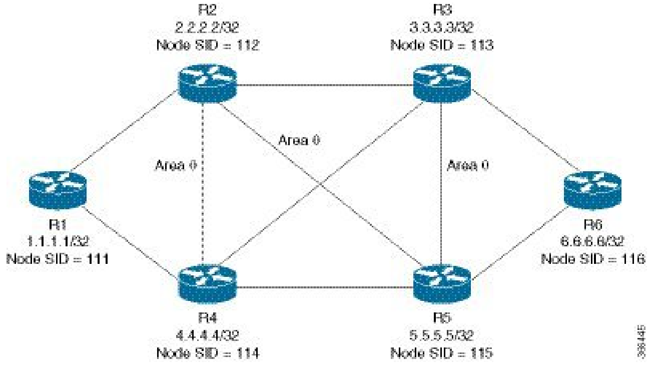
8. If you do not wish a specific interface to be enabled as a repair path, use the candidate disable command.

```
Device(config-if)#ip ospf fast-reroute per-prefix candidate disable
```

Configuring Topology Independent Fast Reroute Tie-breaker

You need to enable segment routing on all the routers with prefix SIDs configured for all the nodes. Use the following topology as a reference to understand the configuration.

Figure 9: Configuration Example



Let us take the device R2 which is protecting the link between R2 and R3. The configuration at R2:

```
router ospf 10
fast-reroute per-prefix enable prefix-priority low
fast-reroute per-prefix ti-lfa
segment-routing mpls
segment-routing area 0 mpls
fast-reroute per-prefix enable prefix-priority low
fast-reroute per-prefix ti-lfa
fast-reroute per-prefix ti-lfa area 0
fast-reroute per-prefix tie-break node-protecting index 60
fast-reroute per-prefix tie-break srlg index 70
mpls traffic-eng router-id Loopback1
mpls traffic-eng area 0

interface GigabitEthernet4   //interface connecting to the router 4
ip address 100.101.4.4 255.255.255.0
ip ospf 10 area 0
ip ospf network point-to-point
srlg gid 10
negotiation auto

interface GigabitEthernet3  //interface connecting to the router 3
ip address 100.101.3.3 255.255.255.0
ip ospf 10 area 0
ip ospf network point-to-point
srlg gid 10
negotiation auto

interface GigabitEthernet5   //interface connecting to the router 2
ip address 100.101.5.5 255.255.255.0
ip ospf 10 area 0
ip ospf network point-to-point
srlg gid 20
negotiation auto



interface loopback2
```

```
ip address 2.2.2.2/32
ip ospf 10 area 0
```

Note In all the other devices, configuration of segment routing and assignment of connected prefix SIDs need to be done.

How Node Protection Works: Using the same topology as an example, let us take the case where you are protecting the link between R2 and R3 and also the prefix which is leant from R6. In that case, let us assume that the primary path for the prefix is via R2-R3. So, our primary path is R2---R3---R6 and we are protecting the link R2---R3.

In this scenario, only link-protection is configured and enabled. When you enable TI-LFA under OSPF process, then you get the following paths provided the cost for all the paths are equal:

R2----R4----R5---R6

R2---R5----R3---R6

R2----R5---R6

If you have only link protection configured, then all the three paths will be chosen and they will share the load amongst them.

If you wish to configure node protection, then the backup would be calculated in such a way that the back up path does not contain the node that you are protecting. In this example, the node R3 in the back up is not required. As a result, only the following two paths would be chosen as the back up paths:

R2----R4----R5---R6

R2----R5---R6

It is possible that R2---R5---R3---R6 have the lesser cost than the above two paths. But since the node protection is configured, only the paths amongst the above two will be considered.

How SRLG Protection Works: SRLG protection further eliminates the back up paths in a such a way that the primary path and the backup does not share the same SRLG ID. Suppose the following back up paths are available:

R2----R4----R5---R6

R2----R5---R6

Then, the SRLG ID of (R2----R4) and (R2----R5) are compared against the primary interface (R2----R3) which is 10. It is noticed that only the interface R2----R5 has different SRLG ID which is 20. So, only the backup path R2---R5---R6 will be chosen.

Verifying Topology Independent Fast Reroute Tunnels

You can use the following command, to check the TI LFA tunnels:

```
Device#show ip ospf fast-reroute ti-lfa tunnels


            OSPF Router with ID (2.2.2.200) (Process ID 10)

                    Area with ID (0)
                    Base Topology (MTID 0)

Tunnel                Interface         Next Hop         Mid/End Point    Label
-------------------------------------------------------------------------------
MPLS-SR-Tunnel2       Et1/1             2.7.0.7          1.1.1.1          16020
MPLS-SR-Tunnel6       Et0/3             2.8.0.0          3.3.3.3          16003
MPLS-SR-Tunnel7       Et1/1             2.7.0.7          1.1.1.1          16020
                                                         5.5.5.5          16005
                                                         3.3.3.3          16003
MPLS-SR-Tunnel5       Et0/3             2.8.0.0          5.5.5.5          16005
MPLS-SR-Tunnel1       Et1/1             2.7.0.7          1.1.1.1          16020
                                                         5.5.5.5          16005
MPLS-SR-Tunnel3       Et1/1             2.7.0.7          6.6.6.6          16006
```

You can use the following command, to check the route in OSPF routing table with primary and repair path:

```
Device#show ip ospf rib 6.6.6.6


            OSPF Router with ID (2.2.2.200) (Process ID 10)


                Base Topology (MTID 0)

OSPF local RIB
Codes: * - Best, > - Installed in global RIB
LSA: type/LSID/originator

*>  6.6.6.6/32, Intra, cost 31, area 0
     SPF Instance 19, age 02:12:11
      contributing LSA: 10/7.0.0.0/6.6.6.6 (area 0)
     SID: 6
     CSTR Local label: 0
     Properties: Sid, LblRegd, SidIndex, N-Flag, TeAnn
     Flags: RIB, HiPrio
      via 2.7.0.7, Ethernet1/1 label 16006
       Flags: RIB
       LSA: 1/6.6.6.6/6.6.6.6
      PostConvrg repair path via 3.3.3.3, MPLS-SR-Tunnel6 label 16006, cost 81, Lbl cnt 1
       Flags: RIB, Repair, PostConvrg, IntfDj, LC Dj
       LSA: 1/6.6.6.6/6.6.6.6
```

You can use the following command, to display the route in the IP routing table:

```
Device#show ip route 6.6.6.6
Routing entry for 6.6.6.6/32
  Known via "ospf 10", distance 110, metric 31, type intra area
  Last update from 2.7.0.7 on Ethernet1/1, 00:25:14 ago
 SR Incoming Label: 16006
  Routing Descriptor Blocks:
  * 2.7.0.7, from 6.6.6.6, 00:25:14 ago, via Ethernet1/1, merge-labels
      Route metric is 31, traffic share count is 1
      MPLS label: 16006
      MPLS Flags: NSF
      Repair Path: 3.3.3.3, via MPLS-SR-Tunnel6
```

Debugging Topology Independent Loop Free Alternate Fast Reroute

You can use the following commands to debug TI-LFA FRR:

```
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>]
```

Examples: OSPFv2 Link-Protection Topology Independent Loop Free Alternate Fast Reroute

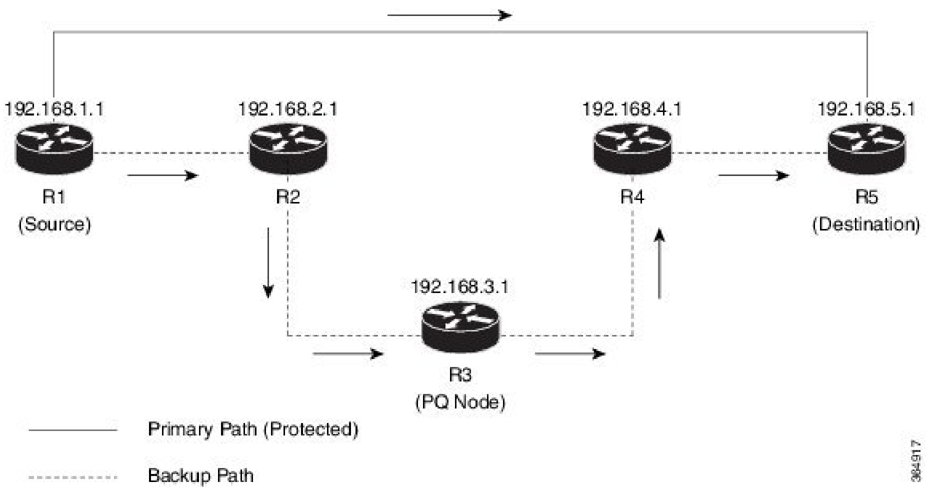
The following are the examples for the OSPFv2 Link-Protection TI-LFA FRR.

Example: Configuring Topology Independent Loop Free Alternate Fast Reroute

This example shows how to configure TI-LFA for segment routing TE tunnels using single or disjoint PQ nodes. The following are the two topologies used:

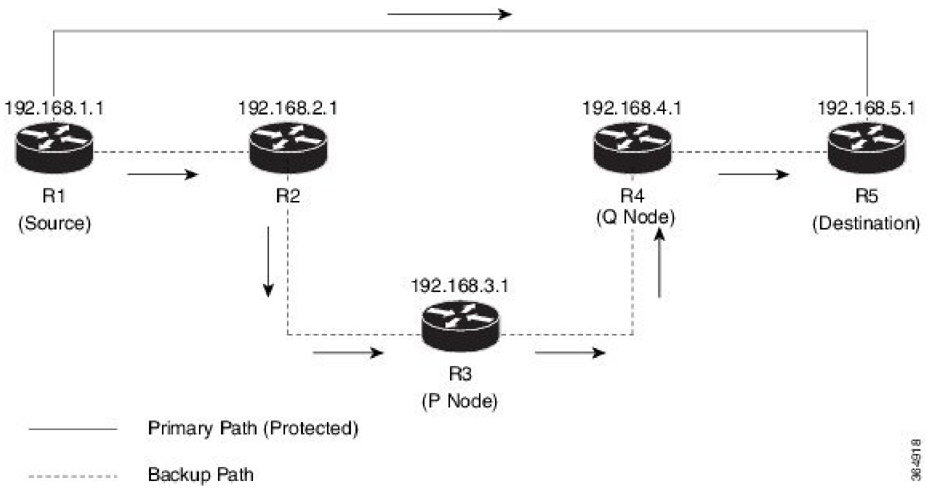
- Topology 1: A single PQ Node and therefore has two SIDs from the source router, R1 through the PQ Node to the destination router, R5.

Figure 10: Topology 1: Single PQ Node



- Topology 2: Disjoint PQ Nodes and therefore consists of three SIDs from the source router R1, through the P Node and the Q Node to the destination router, R5.

Figure 11: Topology 2: Disjoint PQ Nodes



Configure TI-LFA for OSPF on the source router (R1) interface connecting to the destination router (R5).

```
Device(config)# router ospf 10
Device(config-router)# fast-reroute per-prefix enable prefix-priority low
Device(config-router)# fast-reroute per-prefix ti-lfa
Device(config-router)# fast-reroute per-prefix ti-lfa area 0
Device(config-router)# exit
```

Additional References for OSPFv2 Link-protection Topology Independent Loop Free Alternate Fast Reroute

Related Documents

Related Topic	Document Title
Cisco IOS Commands	Cisco IOS Master Command List, All Releases

Feature Information for OSPFv2 Link-protection Topology Independent Loop Free Alternate Fast Reroute

The following table provides release information about the feature or features described in this module. This table lists only the software release that introduced support for a given feature in a given software release train. Unless noted otherwise, subsequent releases of that software release train also support that feature.

Use Cisco Feature Navigator to find information about platform support and Cisco software image support. To access Cisco Feature Navigator, go to www.cisco.com/go/cfn. An account on Cisco.com is not required.

Table 6: Feature Information for OSPFv2 Link-protection Topology Independent Loop Free Alternate Fast Reroute

Feature Name	Releases	Feature Information
OSPFv2 Link-protection Topology Independent Loop Free Alternate Fast Reroute	Cisco IOS XE Everest 16.4.1 Cisco IOS XE Fuji 16.7.1	Topology-Independent Loop-Free Alternate (TI-LFA) uses segment routing to provide link, node, and Shared Risk Link Groups (SRLG) protection in topologies where other fast reroute techniques cannot provide protection. The goal of TI-LFA is to reduce the packet loss that results while routers converge after a topology change due to a link failure. The following commands were introduced or modified: **fast-reroute per-prefix ti-lfa [area <area> [disable]]**, **fast-reroute per-prefix tie-break node-protecting index <index>**, **fast-reroute per-prefix tie-break node-protecting required index <index>**, **fast-reroute per-prefix tie-break srlg index <index>**, **fast-reroute per-prefix tie-break srlg required index <index>**, **ip ospf fast-reroute per-prefix protection disable**, **ip ospf fast-reroute per-prefix candidate disable**, **show ip ospf fast-reroute ti-lfa tunnels**. In Cisco IOS XE Fuji 16.7.1, this feature is supported on Cisco 4000 Series Integrated Service Routers.

CHAPTER 8

Segment Routing Traffic Engineering With OSPF

This chapter describes how Segment Routing traffic engineering can be implemented using OSPF.

Restrictions for Segment Routing Traffic Engineering With OSPF

- Segment Routing Traffic Engineering is supported only on OSPFv2.
- SR-TE is not supported on broadcast interfaces; it is supported only point-to-point interfaces.
- The Cisco ASR routers support only a specific number of labels imposed on an outgoing packet. If the number of labels are greater than the specified number, the SR-TE tunnel creation fails. The Cisco ASR1000 routers support a maximum of 16 labels.
- Only one instance of protocol should be enabled for TE at a given point of time.

Information About Segment Routing Traffic Engineering With OSPF

A Traffic Engineered (TE) tunnel is a container of TE LSP(s) instantiated between the tunnel ingress and the tunnel destination. A TE tunnel may instantiate one or more SR-TE LSP(s) 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 a set of prefix-SID(s) and/or adjacency-SID(s) of nodes and/or links to be traversed by the SR-TE LSP.

The head-end imposes the corresponding MPLS label stack on to 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. OSPF provides TE with the topology and SR related information. SR related information include SRGB/prefix/Adjacency SIDs of all nodes/links with SR enabled in the network.

Benefits of Using Segment Routing Traffic Engineering With OSPF

Segment routing traffic engineering offers a comprehensive support for all useful optimizations and constraints, for example:

- Latency
- Bandwidth
- Disjointness
- Resource avoidance

OSPFv2 provides the following functionalities for SR-TE:

- OSPFv2 provides SR information along with TE topology information to TE module.
- TE uses this information to construct SR TE path/tunnel comprising of one or more segments - with the combination of prefix and/or adjacency segments.
- For the prefixes TE is interested in, OSPF provides first hop resolution to setup the forwarding plane.
- SR TE tunnels are also advertised back into OSPF (like RSVP TE tunnels) for diverting traffic over the SR-TE tunnels.

OSPFv2 Segment Routing Traffic Engineering Functionalities

OSPFv2 perform the following functionalities for SR-TE:

- OSPFv2 provides SR information along with TE topology information to TE module.
- TE uses this information to construct SR TE path/tunnel comprising of one or more segments - with the combination of prefix and/or adjacency segments.
- For the prefixes TE is interested in, OSPF provides first hop resolution to setup the forwarding plane.
- SR TE tunnels are also advertised back into OSPF (like RSVP TE tunnels) for diverting traffic over the SR-TE tunnels.

Protected Adjacency SID

Segment routing creates protected adjacency SID for point to point to point interfaces and broadcast interfaces. It advertises them to the extended link-state advertisement (LSA) along with the unprotected adjacency SID. Protected adjacency SID can have a repair path, but it is not guaranteed to have a repair path.

Traffic Engineering Interfaces

In order to support SR-TE functionality, TE interfaces with various components, and with IGP (OSPF and ISIS) to distribute and receive information on TE topology. For SR-TE support, OSPF needs to additionally provide SR information to TE that it had received through various LSAs, for example,

- Router Information LSA
- Extended Prefix LSA

- Extended Link LSA

TE interfaces distribute information, such as bandwidth resources, constraints, capabilities, and other attributes, associated with the links that are configured for TE. The link information is distributed to other routers using opaque LSAs and is used by TE to create a local topology database. The topology database is a key element in allowing TE to compute a suitable constraint-based path for establishing an LSP. TE also interfaces with the IGP to notify when a TE headend interface can be considered for routing packets.

Unnumbered Support

IS-IS description of an unnumbered link does not contain remote interface ID information. The remote interface ID of an unnumbered link is required to include the unnumbered link as part of the SR-TE tunnel.

Segment Routing Traffic Engineering Support for Forwarding Adjacency

MPLS TE forwarding adjacency feature is supported by OSPF. In this, TE tunnel is considered as a link in the IGP network. TE tunnel interfaces are advertised in the IGP network like any other links. Routers can then use these links to compute the shortest path tree (SPT).

Note This feature is not supported with the SR-TE tunnels.

Segment Routing Traffic Engineering Support for Auto-route Announce

MPLS TE auto-route announce feature is supported by OSPF, that uses TE Tunnel as the first-hop, if the node is reachable via that tunnel. It allows the traffic to the nodes that are downstream to the tail-end of the TE tunnel flows through the tunnel. OSPF supports auto-route over the SR-TE tunnels similar to the MPLS TE tunnels setup using RSVP.

The TE tunnel that instantiates an SR-TE LSP can be Auto-route Announced (AA) into IGP (OSPF and ISIS) as an IGP shortcut. The IGP uses the TE tunnel as next hop and installs routes in RIB for all IP prefixes whose shortest path falls behind the TE tunnel destination. Auto-route announce for of TE tunnels is supported to carry IPV4 prefixes.

Auto-route Announce IP2MPLS

The auto-routeIP2MPLS feature is introduced for SR tunnels to avoid potential packet from looping indefinitely between the SR-TE tunnel headend/ingress and a node that is pointing/routing the packet back to the headend/ingress.

The solution consists in the headend programming in forwarding two sets of path(s) for the prefixes that are mapped over the SR-TE tunnel. The first is the pure IP route for the prefix(es) mapped on the and having the outgoing interface as the tunnel interface. This allows mapping IP traffic directly over the tunnel. The second is the MPLS path for the prefixes mapped on the tunnel. For this the prefix-SID label is programmed with the IGP shortest path outgoing interface(s), that is, non tunnel output interfaces.

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.

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 head-end validates that the specified SR path is compliant with the configured affinity. This necessitates 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.

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 Single Tunnel

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 the head-end or any midpoint traversed node along the SR-TE LSP path.

Load Balancing on Multiple Tunnels

Multiple TE tunnels can be used as next-hop paths for routes to specific IP prefixes either by configuring static route on multiple tunnels, or auto-route announcing multiple parallel tunnels to the same destination. In such cases, the tunnels share the traffic load equally or load balance traffic on multiple parallel tunnels. It is also possible to allow Unequal Load Balance (UELB) with an explicit per tunnel configuration at the tunnel head-end. In this case, the tunnel load-share is passed from MPLS-TE to forwarding plane.

The tunnel load-share feature continues to work for TE tunnels that instantiate the SR-TE LSPs.

SR-TE Tunnel Re-optimization

TE tunnel re-optimization occurs when the head-end determines that there is a more optimal path available than the one currently used. For example, in case of a failure along the SR-TE LSP path, the head-end 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 head-end 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 head-end 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 head-end 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
```

SR-TE With Lockdown Option

The **lockdown** option 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
 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-selection metric igp
 tunnel mpls traffic-eng load-share 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:

IP-FRR Local Repair Protection

On an SR-TE LSP head-end 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 head-end 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 head-end 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.

In the event of a failure of the primary SR-TE LSP, 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 LSP Path Verification

SR-TE tunnel functionality requires that the head-end perform initial verification of the tunnel path as well as the subsequent tracking of the reachability of the tunnel tail-end 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 head-end validates the path of an SR-TE LSP for connectivity against the TE topology. MPLS-TE head-end checks if links corresponding to the adjacency SIDs are connected in the TE topology.

For newly-instantiated SR-TE LSPs, if the head-end 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 head-end detects a discontinuity on any link, the head-end assumes a fault has occurred on that link. In this case, the local repair protection, such as the IP FRR, come 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 head-end 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 head-end 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 head-end 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 head-end. 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 head-end has full visibility over the LSP path, and validates the path to the ultimate LSP destination. However, for inter-area LSPs, the head-end has partial visibility over the LSP path—only up to the first ABR. In this case, the head-end 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 head-end, 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 head-end 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 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 head-end 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 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 head-end 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 commands below. 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
!
```

SR-TE LSP Explicit Null

MPLS-TE tunnel head-end 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 tail-end without a transport label. However, in some cases, it is desirable that the packet arrive at the tail-end with explicit-null label, and in such case, the head-end will impose an explicit-null label at the top of the label stack.

Verbatim Path Support

MPLS TE LSPs usually require that all the nodes in the network are TE aware which means that they have IGP extensions to TE in place. However, some network administrators want the ability to build TE LSPs to traverse nodes that do not support IGP extensions to TE, but that do support RSVP extensions to TE. Verbatim LSPs are helpful when all or some of the intermediate nodes in a network do not support IGP extensions for TE.

When this feature is enabled, the IP explicit path is not checked against the TE topology database. Since the TE topology database is not verified, a Path message with IP explicit path information is routed using the shortest path first (SPF) algorithm for IP routing.

How to Configure Segment Routing Traffic Engineering With OSPF

Perform the following steps to configure Segment Routing Traffic Engineering With OSPF.

Enabling Segment Routing Traffic Engineering With OSPF

OSPF Segment Routing traffic engineering is enabled when the segment-routing is enabled along with mpls traffic engineering. SR-TE support is turned on in an area when you enable SR & MPLS TE in that area.

```
router ospf 10
 router-id 10.10.10.2
 segment-routing mpls
  mpls traffic-eng area 0
```

Configuring Path Option for a TE Tunnel

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

```
Device(config)# interface tunnel 100
Device(config-if)# tunnel mpls traffic-eng path-option 1 explicit name foo segment-routing
Device(config-if)# tunnel mpls traffic-eng path-option 2 dynamic segment-routing
Device(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 will be 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

The following SR-TE explicit path hops are supported:

- IP addresses
- MPLS labels
- Mix of IP addresses and MPLS labels

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

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

Note When using IP unnumbered interfaces, you cannot specify next hop address as an explicit path index. It should be node address or label.

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

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

Configuring 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 head-end validates that the specified SR path is compliant with the configured affinity. This necessitates 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

Perform the following steps to configure 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
```

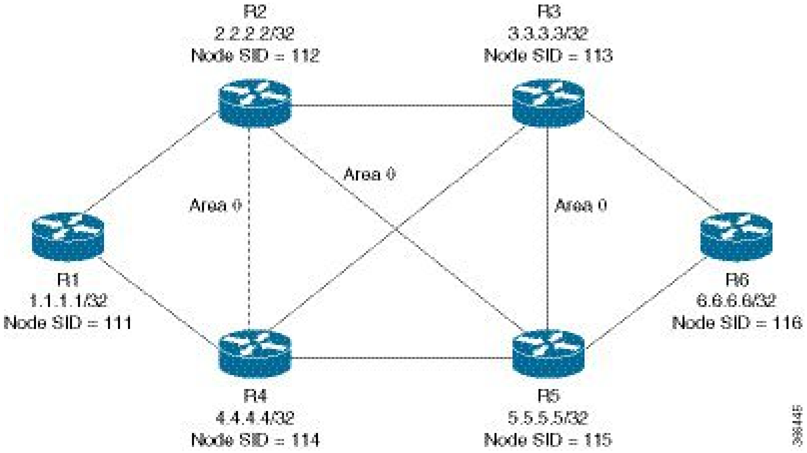
Configuring Segment Routing Traffic Engineering With OSPF

Consider the following inter area and intra area use cases for configuring SR-TE with OSPF:

Configuring Intra Area Tunnel

Consider the following topology to configure intra area tunnel:

Figure 12: Intra Area Tunnel



All the routers are configured in the same area, Area 0.

Configuration at the head end router R1:

```
router ospf 10
fast-reroute per-prefix enable prefix-priority low
fast-reroute per-prefix ti-lfa
segment-routing mpls
mpls traffic-eng router-id Loopback1
mpls traffic-eng area 0

interface GigabitEthernet2   //interface connecting to the router 2
ip address 100.101.1.1 255.255.255.0
ip ospf 10 area 0
ip ospf network point-to-point
negotiation auto
mpls traffic-eng tunnels

interface GigabitEthernet4  //interface connecting to the router 4
ip address 100.101.1.1 255.255.255.0
ip ospf 10 area 0
ip ospf network point-to-point
negotiation auto
mpls traffic-eng tunnels

interface loopback1
ip address 1.1.1.1/32
ip ospf 10 area 0
```

Configuration at the tail-end router R6:

```
router ospf 10
fast-reroute per-prefix enable prefix-priority low
fast-reroute per-prefix ti-lfa
segment-routing mpls
```

```
mpls traffic-eng area 0
mpls traffic-eng router-id Loopback1
interface GigabitEthernet2   //interface connecting to the router 3
ip address 100.101.2.1 255.255.255.0
ip ospf 10 area 0
ip ospf network point-to-point
negotiation auto
mpls traffic-eng tunnels

interface GigabitEthernet4  //interface connecting to the router 5
ip address 100.101.2.1 255.255.255.0
ip ospf 10 area 0
ip ospf network point-to-point
negotiation auto
mpls traffic-eng tunnels

interface loopback1
ip address 6.6.6.6/32
ip ospf 10 area 0
```

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
 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 load-share 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
 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 load-share 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
 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 load-share 10
```

Note In the case of mixed path, IP next-hop cannot be used after using Node SIDs in the path. The following path will not be valid:

```
ip explicit-path name MIXED_PATH enable
next-label 115
next-label 116
next-address 2.2.2.2
```

Dynamic Path SR-TE Tunnel 4

Consider that tunnel 4is based on adjacency SIDs

```
interface Tunnel4
 ip unnumbered Loopback1
 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 load-share 10
end
```

Dynamic Path SR-TE Tunnel 5

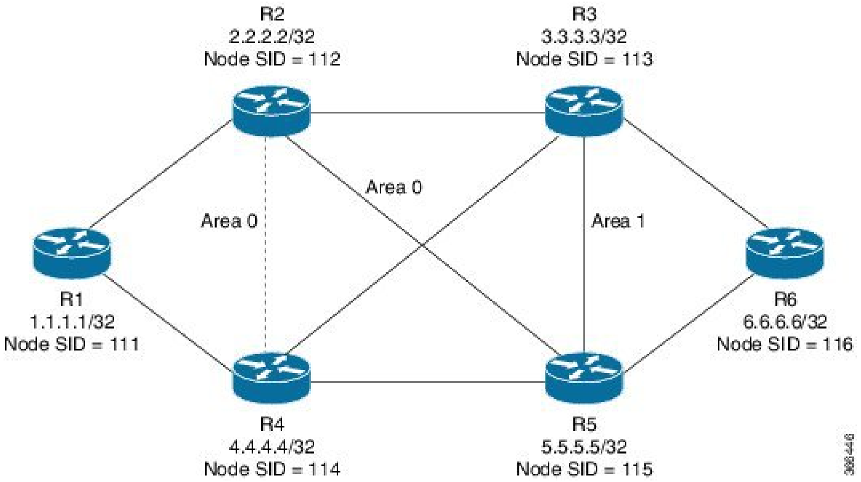
Consider that tunnel 5 is based on Node SIDs

```
 interface Tunnel5
 ip unnumbered Loopback1
 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 load-share 10
```

Configuring Inter Area Tunnel

Consider the following topology to configure inter area tunnel:

Figure 13: Inter Area Tunnel



All the routers are configured in the same area, area 0 except R6 which is configured in area 1.

Configuration at the head end router R1:

```
router ospf 10
fast-reroute per-prefix enable prefix-priority low
fast-reroute per-prefix ti-lfa
segment-routing mpls
mpls traffic-eng router-id Loopback1
mpls traffic-eng area 0

interface GigabitEthernet2   //interface connecting to the router 2
ip address 100.101.1.1 255.255.255.0
ip ospf 10 area 0
ip ospf network point-to-point
negotiation auto
mpls traffic-eng tunnels

interface GigabitEthernet4  //interface connecting to the router 4
ip address 100.101.1.1 255.255.255.0
ip ospf 10 area 0
ip ospf network point-to-point
negotiation auto
mpls traffic-eng tunnels

interface loopback1
ip address 1.1.1.1/32
ip ospf 10 area 0
```

Configuration at the tail-end router R6:

```
router ospf 10
fast-reroute per-prefix enable prefix-priority low
fast-reroute per-prefix ti-lfa
segment-routing mpls
mpls traffic-eng area 1
mpls traffic-eng router-id Loopback1
```

```
interface GigabitEthernet2   //interface connecting to the router 3
ip address 100.101.2.1 255.255.255.0
ip ospf 10 area 1
ip ospf network point-to-point
negotiation auto
mpls traffic-eng tunnels

interface GigabitEthernet4  //interface connecting to the router 5
ip address 100.101.2.1 255.255.255.0
ip ospf 10 area 1
ip ospf network point-to-point
negotiation auto
mpls traffic-eng tunnels

interface loopback1
ip address 6.6.6.6/32
ip ospf 10 area 1
```

Restrictions for Configuring Inter Area Tunnel

The following are the restrictions for configuring inter area tunnel:

- The dynamic option with node and adjacency SID are not supported.
- You can configure inter are tunnel using the explicit path containing only labels and/or IP address and labels.

Note The IP address can be used only be till the Area Border Router (ABR) and after that you need to specify only the labels.

Explicit Path SR-TE Tunnel 1

Consider tunnel 2 is based on node SIDs.

```
ip explicit-path name IA_PATH
next-label 114
next-label 115
next-label 116
!
interface Tunnel2
ip unnumbered Loopback1
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 load-share 10
end
```

Explicit Path SR-TE Tunnel 2

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
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 load-share 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 (ospf 10  area 0)
    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 (ospf 10  area 0)
    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 (ospf 10  area 0)
    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 (ospf 10  area 0)
    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 (ospf 10  area 0)
    Segment0[Node]: 6.6.6.6, Label: 116
```

Additional References for Segment Routing Traffic Engineering With OSPF

Related Documents

Related Topic	Document Title
Cisco IOS commands	Cisco IOS Master Commands List, All Releases

Feature Information for Segment Routing Traffic Engineering With OSPF

The following table provides release information about the feature or features described in this module. This table lists only the software release that introduced support for a given feature in a given software release train. Unless noted otherwise, subsequent releases of that software release train also support that feature.

Use Cisco Feature Navigator to find information about platform support and Cisco software image support. To access Cisco Feature Navigator, go to www.cisco.com/go/cfn. An account on Cisco.com is not required.

Table 7: Feature Information for Segment Routing Traffic Engineering With OSPF

Feature Name	Releases	Feature Information
Segment Routing Traffic Engineering With OSPF	Cisco IOS XE Release 3.17S Cisco IOS XE Fuji 16.7.1	A Traffic Engineered (TE) tunnel is a container of TE LSP(s) instantiated between the tunnel ingress and the tunnel destination. A TE tunnel may instantiate one or more SR-TE LSP(s) that are associated with the same tunnel. The following commands were added or modified: **show mpls traffic-eng tunnels**, **tunnel mpls traffic-eng path-option 10 dynamic segment-routing**, **tunnel mpls traffic-eng path-option 10 segment-routing**, **tunnel mpls traffic-eng path-option 10 explicit name MIXED_PATH segment-routingtunnel mpls traffic-eng path-option 10 explicit name IP_PATH1 segment-routingtunnel mpls traffic-eng path-option 10 explicit name NODE_PATH segment-routing**. In Cisco IOS XE Fuji 16.7.1, this feature is supported on Cisco 4000 Series Integrated Service Routers.

CHAPTER 9

BGP Dynamic Segment Routing Traffic Engineering

Border Gateway Protocol (BGP) has become a popular choice as a routing protocol in Data Center (DC) network. The ability to setup Segment Routing-Traffic Engineering (SR-TE) path initiated by BGP simplifies DC network operation.

Restrictions for Segment Routing –Traffic-Engineering Dynamic BGP

- For Anycast SID support to work BGP-TE should be configured with the prepend feature.
- In the case of BGP Dynamic SR-TE if SR-TE fails, forwarding gets broken.

Information About Segment Routing –Traffic-Engineering Dynamic BGP

In BGP dynamic SR-TE, the label Switched Path (LSP) is enabled on demand when defined criteria and policies are met and that is the key difference between manually enabled SR-TE and BGP dynamic SR-TE. Policies, for example, low latency path, minimum cost path, and so on are carried via BGP and matches on a given customer prefix. SR-TE tunnel used for L3VPN or Virtual Private LAN Services (VPLS) using BGP for auto-discovery and signaling is referred to as BGP-TE Dynamic.

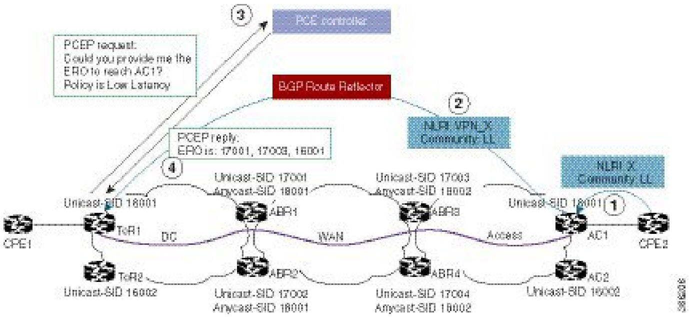
BGP SR-TE dynamic assumes the on-demand auto-tunnel resides in single IGP domain. In this case path computation is done via IGP. SR-TE auto-tunnel created based on the request from BGP is a dynamic SR-TE tunnel. In other words, tunnel path information, or label stack, is computed based on the BGP next-hop and TE attribute configuration. BGP dynamic SR-TE functions to trigger an On-demand LSP (auto-tunnel). The functions include:

- Tag customer prefixes (IPv4 or L3VPN VRF) using communities (community list) via route map configuration.
- Associate each community with a TE attribute-set or profile.

SR-TE profile is locally configured in attribute-set to define certain SR-TE parameters, for example, latency, disjoint path and so on. Once the BGP customer prefixes are mapped to an SR-TE-profile, a tunnel is dynamically created (auto-tunnel or On demand Label Switched Path (LSP)) using the parameters defined in the attribute-set, for each specified BGP next-hop and attribute-set pair associated with the prefixes. A binding SID is associated with each SR-TE auto-tunnel and passed to BGP. The binding SID or binding label is installed into Routing Information Base (RIB) and Forwarding Information Base (FIB). FIB resolves BGP path via the binding SID or binding label, which forwards over the On demand SR-TE auto-tunnel. The binding-SID is also used to steer the customer traffic over the SR-TE LSP.

It must be noted that BGP only carries the SR-TE policy in this case, while path computation is done via IGP in a single IGP domain. In a single IGP domain the headend node has full visibility of the end to end path and the topology engineering database (Traffic Engineering Database or TED). Also it is assumed with BGP Dynamic SR-TE that all the nodes reside within single AS and single IGP domain.

Figure 14: BGP-TE Dynamic Workflow



The above figure depicts the workflow for BGP-TE dynamic using multiple routing domains use case:

1. Customer premise equipment 2 (CPE) sends BGP update for Prefix-X and adds LL community, for example, 100:333.
2. AC1 announces a VPN route for prefix X with LL community.
3. After receiving BGP update of the VPN route matching community LL, ToR1 sends a request to PCE controller for LSP path towards AC1 with low latency TE policy.
4. Path calculation element (PCE) controller replies with a label stack, for example, 17003, 1600.
5. ToR1 creates SR-TE auto-tunnel and installs the route for Prefix-X in VRF of this VPN.

TE Label Switched Path Attribute-Set

TE-LSP attribute-set is used to configure the properties of a LSP. It describes TE profile or policy such as bandwidth, affinities inclusion and exclusion, links/nodes/SRLG inclusion and exclusion, metrics, path disjoint degree and group, and so on that are used to create an auto-tunnel.

How to Configure TE Label Switched Path Attribute-Set

Configuring TE Label Switched Path Attribute-Set

You can use the command **mpls traffic-eng lsp attribute** <*name*> to configure TE-LSP attribute. The following options are available:

```
Mpls traffic-eng lsp attribute name
   affinity      Specify attribute flags for links comprising LSP
   lockdown      Lockdown the LSP--disable reoptimization
   priority      Specify LSP priority
```

TE-LSP attribute command can be extended to support configuration for the two options **pce** and **path-selection**. It can be configured as following:

```
mpls traffic-eng lsp attribute name <test>
    path-selection
       metric <te/igp>
       invalidation <time-out> <drop/tear>
       segment-routing adjacency <protected/unprotected>
```

- If pce option is set in the TE attribute the dynamic path is calculated by PCE. Otherwise, the path is calculated locally by TE PCALC (path-calculation) entity. In the later case, IGP has to be configured and the BGP next-hop has to be both advertised by IGP and reachable from the local node over an IGP route.
- The option path-selection metric indicates whether the path calculation is based on TE metrics or IGP metrics. If this option is not configured the global value configured under mpls traffic-eng path-selection metric is used.
- The option **path-selection invalidation** configures the behavior of how an LSP reacts to soft failure from network. When an LSP path has a protected path from IGP against a link or node failure, the failure to the link or node is considered as soft failure.
- The option **path-selection segment-routing adjacency** indicates whether to choose an adjacency-SID with or without IGP protection when calculating LSP label stack.
- The option **pce disjoint-path** indicates the tunnel LSP is a member of disjoint-path group. Any LSPs within the same disjoint-path group do not traverse the same resources, such as links, nodes, or SRLG, in its path. This is used to create two or more tunnel LSPs with disjoint paths.

For BGP-TE Dynamic, a TE attribute name is associated with a BGP route-map set extension as following:

```
route-map <name>
    match community <name>
         set attribute-set  <name>
```

BGP uses the **attribute-set** <*name*> string together with its BGP next-hop to request a SR-TE auto-tunnel.

Additional References for BGP Dynamic Segment Routing Traffic Engineering

Related Documents

Related Topic	Document Title
Cisco IOS Commands	Cisco IOS Master Command List, All Releases

Feature Information for BGP Dynamic Segment Routing Traffic Engineering

The following table provides release information about the feature or features described in this module. This table lists only the software release that introduced support for a given feature in a given software release train. Unless noted otherwise, subsequent releases of that software release train also support that feature.

Use Cisco Feature Navigator to find information about platform support and Cisco software image support. To access Cisco Feature Navigator, go to www.cisco.com/go/cfn. An account on Cisco.com is not required.

Table 8: Feature Information for BGP Dynamic Segment Routing Traffic Engineering

Feature Name	Releases	Feature Information
BGP Dynamic Segment Routing Traffic Engineering	Cisco IOS XE Everest 16.5.1b	In BGP dynamic SR-TE, the label Switched Path (LSP) is enabled on demand when defined criteria and policies are met. The following commands were introduced or modified: **mpls traffic-eng lsp attribute** *name*

CHAPTER 10

Segment Routing On Demand Next Hop for L3/L3VPN

When redistributing routing information across domains, provisioning of multi-domain services (L2VPN & L3VPN) has its own complexity and scalability issues. 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 replied multi-domain LSP for the duration of the service into the local forwarding information base (FIB).

Restrictions for Segment Routing On Demand Next Hop for L3/L3VPN

- On Demand Next Hop (ODN) anycast SID is not supported.
- ODN for IPv6 is not supported.
- SR ODN tunnel is not supported with BGP Nonstop Routing (NSR). It is only supported with BGP Nonstop Forwarding (NSF).

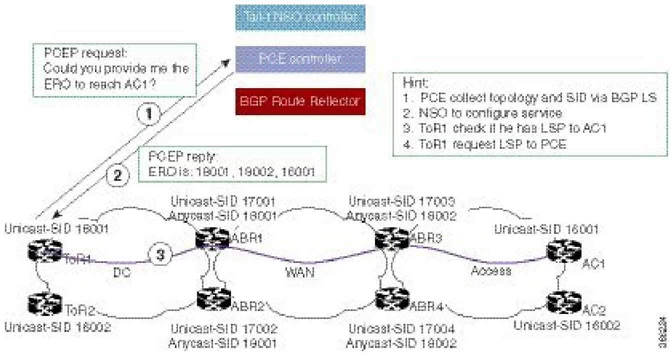
  To enable BGP NSF, use the following command:

  ```
  bgp grace-full restart
  neighbor 10.0.0.2 ha-mode graceful-restart
  ```

Information About Segment Routing On Demand Next Hop for L3/L3VPN

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 below 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 15: ODN Operation



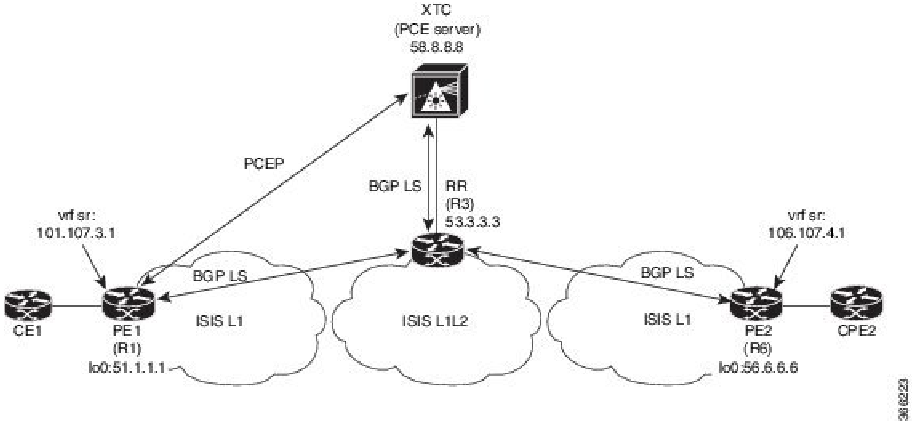
1. PCE controller collects topology and SIDs information via BGP Link State (BGP-LS). For more information on BGP-LS, refer BGP Link-State.
2. If NSO controller is enable, it configures L3VPN VRF or IPv4 prefixes and requests are sent to ToR1 and AC1.
3. ToR1 and AC1 checks if a LSP towards each other 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 via BGP.
4. PCE controller computes the path and replies with a label stack (18001, 18002, 16001, example in ToR1).
5. ToR1 and AC1 create a SR-TE auto-tunnel and reply back to the NSO controller indicating that the LSP for VRF or IPv4 is up and operational.

How to Configure Segment Routing On Demand Next Hop for L3/L3VPN

Configuring Segment Routing On Demand Next Hop for L3/L3VPN

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

Figure 16: ODN Auto-Tunnel Setup



1. Configure the router (R6 tail end) with VRF interface.

```
interface GigabitEthernet0/2/2
 vrf forwarding sr
 ip address 10.0.0.1 255.0.0.0
 negotiation auto

interface Loopback0
 ip address 192.168.0.1 255.255.0.0
 ip router isis 1
```

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

```
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
```

3. Enable BGP on R6 (tail end) and R1 (head end) to advertise and receive VRF SR prefix and match on community set on R6 (tail end).

```
 router bgp 100
 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
 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/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 BGP-LS session is properly established with RR.

```
Device# sh 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# sh 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)
```

Verifying Segment Routing On Demand Next Hop for L3/L3VPN

The ODN verifications are based on L3VPN VRF prefixes.

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

```
Device# sh 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# sh 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
```

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

```
Device# sh 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# sh 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        10.0.0.7/32 is directly connected, GigabitEthernet0/3/1
      10.0.0.8/24 is subnetted, 1 subnets
B        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# sh 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# sh 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# sh ip cef label-table
Label               Next Hop             Interface
0                    no route
865                  attached             Tunnel2000

     Device# sh 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# sh 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 from PCE server: 10.0.0.3--□ via PCE server
      PCE metric: 30, type: IGP
    Reported paths:
      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
  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# sh 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# sh pce lsp summ

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# sh pce lsp det
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. Shutdown the interface that is connected to VRF SR so that the prefix is no longer advertised by MP-BGP.

```
Device# int gig0/2/2
Device(config-if)#shut
```

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

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

Additional References for Segment Routing On Demand Next Hop for L3/L3VPN

Related Documents

Related Topic	Document Title
Cisco IOS Commands	Cisco IOS Master Command List, All Releases

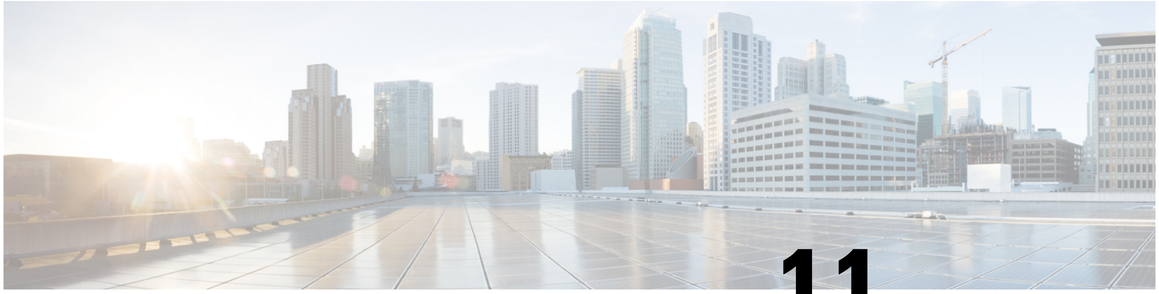
Feature Information for Segment Routing On Demand Next Hop for L3/L3VPN

The following table provides release information about the feature or features described in this module. This table lists only the software release that introduced support for a given feature in a given software release train. Unless noted otherwise, subsequent releases of that software release train also support that feature.

Use Cisco Feature Navigator to find information about platform support and Cisco software image support. To access Cisco Feature Navigator, go to www.cisco.com/go/cfn. An account on Cisco.com is not required.

Table 9: Feature Information for Segment Routing On Demand Next Hop for L3/L3VPN

Feature Name	Releases	Feature Information
Segment Routing On Demand Next Hop for L3/L3VPN	Cisco IOS XE Everest 16.5.1b	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. The following commands were introduced or modified: **route-map BGP_TE_MAP permit**, **mpls traffic-eng tunnels**, **sh bgp li li summary**, **sh pce client peer**, **sh pce ipv4 peer**, **sh ip route vrf sr**, **sh ip bgp vpnv4 vrf sr**, **sh ip cef label-table**, **sh mpls traffic-eng tunnels**, **sh pce client lsp brief**, **sh pce lsp summ**, **sh pce lsp det**, **routing-default-optimize**



CHAPTER 11

Routing Information Base Support

The Routing Information Base (RIB) enhancement supports route redistribution and on-demand nexthop requirements.

Routing Information Base Support for Route Redistribution

Effective with Cisco IOS XE Everest 16.5.1, a requirement to redistribute labels associated with prefixes is introduced. To support redistribution requirements, the storage of local label per prefix is supported in RIB.

The local label is stored instead of the SID to ease use with different protocols which may use different SRGBs. The SID assigned by the destination protocol may not be the same as the SID associated with the source protocol.

The prefix reachability advertisement or an SRMS advertisement is the source of the SID. In SRMS advertisement, the destination protocols for redistribution do not advertise the SID in their prefix reachability advertisements, as this alters conflict resolution by indicating on other network nodes that the source of the advertisement was not from SRMS.

OSPF Node SID Redistribution Support

Effective Cisco IOS XE 16.7.1, when OSPF receives the redistributed prefixes from other IGPs and vice versa the prefix segment identifiers (SIDs) are also advertised which was not the case earlier. You needed to have the BGP LS (or) segment routing mapping server (SRMS) support to learn the SIDs across the IGP domains.

When the user enable redistribution under OSPF the prefix SID entries associated with the prefix entries are provided to OSPF. This gets advertised by OSPF to all its neighbor. The way OSPF advertises varies depending upon the role of OSPF in the network.

Information About OSPF Node SID Redistribution Support

NSSA ASBR

When you enable **redistribute ISIS** *instance* **ip** under OSPF which is Not-So-Stubby Area autonomous system boundary router (NSSA ASBR), it gets all the prefixes from IP routing information base (RIB) which are learnt by IS-IS along with the SID entries. OSPF generates Extended Prefix LSA (EPL) with the scope as area and the route type as RTYPE_NSSA1 or RTYPE_NSSA2 for the prefixes and advertises to all its neighbors. Similarly, when the redistribution is un-configured (or) when the prefixes become unavailable OSPF withdraws the EPL. When the redistributed route is a non-connected route then the OSPF sets the No-PHP flag but explicit NULL flag is not set. However, when the redistributed route is a connected route then OSPF sets the explicit NULL and No-PHP flag according to the configuration done in the SR policy.

When NSSA ABR receives the EPL, the ABR translates the LSA into opaque AS EPL and floods it to all its neighbors.

When a NSSA router which is neither ABR nor ASBR receives the EPL, it learns the prefix along with the SID entries and floods it to all its neighbors in the same area.

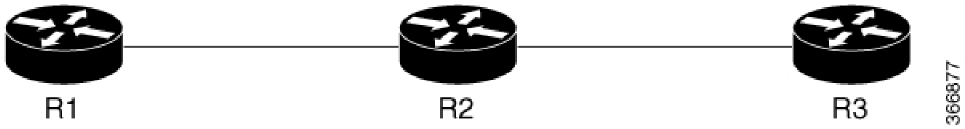
non-NSSA ASBR

When the user enabled **redistribute ISIS** *instance* **ip** under OSPF which is regular ASBR router, it gets all the prefixes from IP RIB which are learnt by IS-IS along with the SID entries. OSPF generates EPL with the scope as autonomous system (AS) and the route type as RTYPE_EXTERN1 or RTYPE_EXTERN2 for the prefixes and advertises to all its neighbors. Similarly when the redistribution is unconfigured (or) when the prefixes become unavailable, OSPF withdraws the EPL again with AS-Scope. When the redistributed route is a non-connected route then the OSPF sets the No-PHP flag but explicit NULL flag is not set. However, when the redistributed route is a connected route then OSPF sets the explicit NULL and No-PHP flag according to the configuration done in the SR policy. When a router receives the EPL with AS scope, it learns the prefix along with the SID entry and floods it to all its neighbors in all areas.

Redistributing Prefix

When IS-IS is enabled for redistribution of OSPF routes the prefixes are given along with the SID information so that the prefixes reach to other domain with the SID values. Refer to the below topology to understand the OSPF prefixes redistribution to the other domains:

Figure 17: OSPF Prefix Redistribution



R1 and R2 are enabled for OSPF. R2 and R3 are enabled for IS-IS. Both IS-IS and OSPF are enabled for Segment Routing. In R2, both IS-IS and OSPF are configured. Prefixes configured are:

1. 1.1.1/32 in R1 (enabled for OSPF with SID 1)
2. 2.2.2/32 in R2 (enabled for OSPF with SID 2)
3. 3.3.3/32 in R3 (enabled for ISIS SID 3)

When you enable SID redistribution in R2, then the prefix 3.3.3.3/32is redistributed to R1. So, R1 knows the SID to reach the prefix R3.

```
conf t
router isis 10
 net 49.0001.0000.0000.0001.00
 metric-style wide
distribute link-state
 segment-routing mpls
router ospf 10
router-id 2.2.2.2
segment-routing mpls
distribute link-state
```

To enable redistribution of ISIS into OSPF routes:

```
conf t
router ospf 10
redistribute isis 10 ip
```

Verifying OSPF Node SID Redistribution

Use the **show ip ospf rib redistribution detail** command to verify if OSPF is redistributing the prefixes from IS-IS.

```
Device# show ip ospf rib redistribution  detail
OSPF Router with ID (2.2.2.2) (Process ID 10)

                Base Topology (MTID 0)

OSPF Redistribution
3.3.3.3/32, type 2, metric 20, tag 0, from IS-IS Router
  Attributes 0x1000000, event 1, PDB Index 4, PDB Mask 0x0
  Source route metric 20, tag 0
  SID 1003, SID Flags NP-bit, EPX Flags None
   via 7.9.0.9, Ethernet0/0
```

Use the **show ip ospf segment-routing local-prefix** command to verify if the SID entries are advertised to its neighbor.

```
Device# show ip ospf segment-routing local-prefix

            OSPF Router with ID (2.2.2.2) (Process ID 10)
Area 0:
 Prefix:            Sid:   Index:          Type:      Source:
2.2.2.2/32         2      0.0.0.0         Intra      Loopback0
AS external:
Prefix:            Sid:   Index:          Type:      Source:
 3.3.3.3/32        3     0.0.0.1         External   Redist
```

Use the **show ip ospf segment-routing sid-database** command to verify if the SIDs are received.

```
Device# show ip ospf segment-routing sid-database

            OSPF Router with ID (1.1.1.1) (Process ID 10)
OSPF Segment Routing SIDs

Codes: L - local, N - label not programmed,
       M - mapping-server

SID             Prefix              Adv-Rtr-Id       Area-Id  Type
--------------  ------------------  ---------------  -------  ----------
```

```
1               1.1.1.1/32          1.1.1.1         0        Intra
2               2.2.2.2/32          2.2.2.2         0        Intra
3               3.3.3.3/32          2.2.2.2         -        External
```

Use the **show ip route 3.3.3.3** command to verify if the IP routing entry is configured for the redistributed route.

```
Device# show ip route 3.3.3.3
Routing entry for 3.3.3.3/32
  Known via "ospf 10", distance 110, metric 20, type extern 2, forward metric 20
  Last update from 1.2.0.2 on Ethernet0/1, 00:00:01 ago
 SR Incoming Label: 16003
  Routing Descriptor Blocks:
  * 1.3.1.3, from 2.2.2.2, 00:00:01 ago, via Ethernet1/1, merge-labels
      Route metric is 20, traffic share count is 1
      MPLS label: 16003
      MPLS Flags: NSF
```

Routing Information Base Support for On-Demand Next Hop

For On-Demand Next Hop (ODN) requirements, RIB supports a next hop called binding label which is provided by the supporting routing protocol (BGP). The binding label is used by the FIB to dynamically resolve the next hop.

The route producer installs a local binding label which identifies the ODN tunnel path associated with the next hop. The labeled traffic is sent via the tunnel and the label is distinct from the existing outlabel.

The following is the sample output of **show ip route** command where each next hop is updated to show the binding label.

```
Device# show ip route 10.10.10.2

Routing entry for 10.10.10.2/32
  Known via "isis", distance 115, metric 10, type level-1
  Redistributing via isis
  Last update from 200.200.200.2 on Ethernet0/0, 00:00:14 ago
  Incoming Label: 16100
  Routing Descriptor Blocks:
  * 200.200.200.2, from 10.10.10.2, 00:00:14 ago, via Ethernet0/0
      Route metric is 10, traffic share count is 1
      * Binding Label 4020, from 2.2.2.2, 00:00:14 ago,
      Route metric is 10, traffic share count is 1
```

Note The incoming labels are seen only after the SID redistribution is enabled.

Additional References for Routing Information Base Support

Related Documents

Related Topic	Document Title
Cisco IOS Commands	Cisco IOS Master Command List, All Releases

Feature Information for Routing Information Base Support

The following table provides release information about the feature or features described in this module. This table lists only the software release that introduced support for a given feature in a given software release train. Unless noted otherwise, subsequent releases of that software release train also support that feature.

Use Cisco Feature Navigator to find information about platform support and Cisco software image support. To access Cisco Feature Navigator, go to www.cisco.com/go/cfn. An account on Cisco.com is not required.

Table 10: Feature Information for Routing Information Base Support

Feature Name	Releases	Feature Information
Routing Information Base Support	Cisco IOS XE Everest 16.5.1b	The Routing Information Base (RIB) enhancement supports route redistribution and On-Demand Nexthop requirements. No new commands were added or modified.
OSPF Node SID Redistribution Support	Cisco IOS XE Fuji 16.7.1	Effective the Cisco IOS XE Fuji 16.7.1 release, when OSPF receives the redistributed prefixes from other IGPs and vice versa the prefix segment identifiers (SIDs) are also advertised which was not the case earlier. You need to have the BGP LS (or) segment routing mapping server (SRMS) support to learn the SIDs across the IGP domains. The following commands were added or modified for this feature: **show ip ospf rib redistribution detail**, **show ip ospf segment-routing local-prefix**, **show ip ospf segment-routing sid-database**, **show ip route 3.3.3.3**.

CHAPTER 12

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. 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 will be used to transport the VPN services between the Metro access rings.

Restrictions for SR-TE On Demand LSP

- Segment-Routing auto tunnel static route does not support ECMP.
- Metrics for IP explicit path and administrtive 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 will fail 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

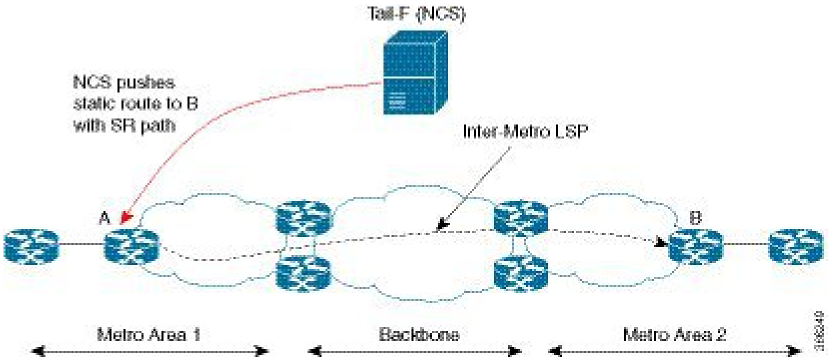
The SR TE On demand LSP feature provides the ability to connect Metro access rings via a static route to the destination.

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 will be used to transport the VPN services between the Metro access rings.

Figure 18: 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
ip route 172.16.0.1 255.240.0.0 segment-routing mpls path name path1
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
```

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.

```
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 summmary

Signalling Summary:
    LSP Tunnels Process:            running
    Passive LSP Listener:           running
    RSVP Process:                   running
    Forwarding:                     enabled
    auto-tunnel:
        p2p    Enabled  (1), id-range:1000-2000
    Periodic reoptimization:        every 3600 seconds, next in 1265 seconds
    Periodic FRR Promotion:         Not Running
    Periodic auto-bw collection:    every 300 seconds, next in 66 seconds
    SR tunnel max label push:       13 labels
    P2P:
      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:   0 established, 0 proceeding, 0 associated, 0 standby
    LSPs Mid:    0 established, 0 proceeding, 0 associated, 0 standby
    LSPs Tail:   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 mpls traffic-eng tunnel auto-tunnel** only displays TE service auto tunnel.

```
Device# show mpls traffic-eng tunnel auto-tunnel detail

 P2P TUNNELS/LSPs:

Name: R1_t1000                            (Tunnel1000) Destination: 0.0.0.0 Ifhandle: 0x17
 (auto-tunnel for ipv4static)
  Status:
    Admin: up         Oper: up     Path: valid       Signalling: connected
    path option 1, (SEGMENT-ROUTING) type explicit (verbatim) path202 (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: 33 days, 20 hours, 29 minutes
      Time since path change: 10 days, 19 hours, 45 minutes
      Number of LSP IDs (Tun_Instances) used: 1646
    Current LSP: [ID: 1646]
      Uptime: 10 days, 19 hours, 45 minutes
    Prior LSP: [ID: 1645]
      ID: path option unknown
      Removal Trigger: signalling shutdown
  Tun_Instance: 1646
  Segment-Routing Path Info (IGP information is not used)
    Segment0[First Hop]: 0.0.0.0, Label: 16002
    Segment1[ - ]: Label: 16006
```

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/up
R1_t3                            66.66.66.66      -         -         up/up
R1_t10                           66.66.66.66      -         -         up/up
SBFD tunnel                      33.33.33.33      -         -         up/up
SBFD Session configured: 1       SBFD sessions UP: 1
```

Additional References for SR-TE On Demand LSP

Related Documents

Related Topic	Document Title
Cisco IOS Commands	Cisco IOS Master Command List, All Releases

Feature Information for SR-TE On Demand LSP

The following table provides release information about the feature or features described in this module. This table lists only the software release that introduced support for a given feature in a given software release train. Unless noted otherwise, subsequent releases of that software release train also support that feature.

Use Cisco Feature Navigator to find information about platform support and Cisco software image support. To access Cisco Feature Navigator, go to www.cisco.com/go/cfn. An account on Cisco.com is not required.

Table 11: Feature Information for SR-TE On Demand LSP

Feature Name	Releases	Feature Information
SR-TE On Demand LSP	Cisco IOS XE Everest 16.5.1b	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 will be used to transport the VPN services between the Metro access rings. The following command was modified: **mpls traffic-eng auto-tunnel**.

CHAPTER 13

Segment Routing MPLS OAM Support

Segment Routing Operations, Administration, and Maintenance (OAM) helps service providers to monitor label-switched paths (LSPs) and quickly isolate forwarding problems to assist with fault detection and troubleshooting in the network. The Segment Routing OAM feature provides support for Nil-FEC (forwarding equivalence classes) LSP Ping and Traceroute, IGP prefix SID FEC type, and partially IGP adjacency-SID FEC type for SR-TE functionality.

Restrictions for Segment Routing OAM MPLS Support

- Ping and traceroute are unsupported with SR-TE static auto tunnel, BGP Dynamic TE, and on-demand next hop auto tunnels.

- Strict-SID option is not supported by the path installed by OSPF.

- MPLS traceroute does not support popping of two explicit null labels in one node.

- Rerouting the path to IP over MPLS segment without using Layer3 VPN is not supported due to IP routing destination not being a MPLS FEC.

Information About Segment Routing MPLS OAM Support

Segment Routing OAM Support

The Nil-FEC LSP ping and traceroute operations are extensions of regular MPLS ping and traceroute . Nil-FEC LSP Ping/Trace functionality support Segment Routing and MPLS Static. It also act as an additional diagnostic tool for all other LSP types. This feature allows operators to test any label stack 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.

Benefits of Segment Routing OAM Support

- The feature enables the MPLS OAM functionality in the Segment Routing Network where the traffic is engineering via SR-TE tunnels or native SR forwarding.
- In traditional MPLS networks, source node chooses the path based on hop by hop signaling protocols such as LDP or RSVP-TE. In Segment Routing Networks, the path is specified by set of segments which are advertised by the IGP protocols (currently OSPF and ISIS).
- As the volume of services offered using SR increase, it is important that the operator essentially is able to do the connectivity verification and the fault isolation in the SR architecture.
- The segment assignment is not based on hop by hop protocols as in traditional MPLS network, any broken transit node could lead in traffic blackholing, which could lead to undesired behavior.
- Both SR and SR-TE supports load balancing, it is important to trace all the ECMP paths available between source and target routers. The features offers the multipath traceroute support for both TE and native SR paths.
- The following are the main benefits of Segment Routing-OAM Support:
    - **Operations**: Network monitoring and fault management.
    - **Administration**: Network discovery and planning.
    - **Maintenance**: Corrective and preventive activities, minimize occurrences and impact of failures.

Segment Routing MPLS Ping

MPLS ping and traceroute are extendable by design. You can add SR support by defining new FECs and/or additional verification procedures. MPLS ping verifies MPLS data path and performs the following:

- Encapsulates echo request packet in MPLS labels.

- Measures coarse round trip time.
- Measures coarse round trip delay.

Segment Routing MPLS Traceroute

MPLS ping and traceroute are extendable by design. You can add SR support by defining new forwarding equivalence classes (FECs) and/or additional verification procedures. MPLS traceroute verifies forwarding and control plane at each hop of the LSP to isolate faults. Traceroute sends MPLS echo requests with monotonically increasing time-to-live (TTL), starting with TTL of 1. Upon TTL expiry, transit node processes the request in software and verifies if it has an LSP to the target FEC and intended transit node. The transit node sends echo reply containing return code specifying the result of above verification and label stack to reach the next-hop, as well as ID of the next-hop towards destination, if verification is successful. Originator processes echo reply to build the next echo request containing TTL+1. Process is repeated until the destination replies that it is the egress for the FEC.

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

Using 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 nil-fec labels <comma separated labels> output interface <tx-interface> nexthop
<nexthop ip addr>
[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-disposition ra-label]
{dsmap | ddmap [l2ecmp]} [hashkey {none | {ipv4 | ipv4-label-set {bitmap <bitmap_size>}}]
```

For more information, refer ping mpls.

Using LSP Traceroute for Nil FEC Target

```
trace mpls nil-fec labels <comma separated labels> output interface <tx-interface>} nexthop
 <nexthop ip addr>
[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]]
[flags {fec | ttl}]
[hashkey ipv4 | ipv4-label-set {bitmap <bitmap_size>}]
```

For more information, refer to the traceroute mpls.

Example for 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#sh 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

Device#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


Device#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
```

Path Validation in Segment Routing Network

The MPLS OAM mechanisms help with fault detection and isolation for a MPLS data-plane path by the use of various target FEC stack sub-TLVs that are carried in MPLS echo request packets and used by the responder for FEC validation. While it is obvious that new sub-TLVs need to be assigned for segment routing, the unique nature of the segment routing architecture raises the need for additional operational considerations for path validation.

The forwarding semantic of Adjacency Segment ID is to pop the Segment ID and send the packet to a specific neighbor over a specific link. A malfunctioning node may forward packets using Adjacency Segment ID to an incorrect neighbor or over an incorrect link. The exposed Segment ID (of an incorrectly forwarded Adjacency Segment ID) might still allow such packet to reach the intended destination, although the intended strict traversal has been broken. MPLS traceroute may help with detecting such a deviation.

The format of the following Segment ID sub-TLVs follows the philosophy of Target FEC Stack TLV carrying FECs corresponding to each label in the label stack. This allows LSP ping/traceroute operations to function when Target FEC Stack TLV contains more FECs than received label stack at responder nodes. Three new sub-TLVs are defined for Target FEC Stack TLVs (Type 1), Reverse-Path Target FEC Stack TLV (Type 16) and Reply Path TLV (Type 21).

sub-Type	Value Field
34	IPv4 IGP-Prefix Segment ID
35	IPv6 IGP-Prefix Segment ID
36	IGP-Adjacency Segment ID

MPLS Ping and Traceroute for IGP Prefix-SID FEC Type

MPLS ping and traceroute operations for prefix SID are supported for various IGP scenarios, for example:

- Within an IS-IS level or OSPF area
- Across IS-IS levels or OSPF areas
- Route redistribution from IS-IS to OSPF and from OSPF to IS-IS

The MPLS LSP Ping feature is used to check the connectivity between ingress Label Switch Routers (LSRs) and egress LSRs along an LSP. MPLS LSP ping uses MPLS echo request and reply messages, similar to Internet Control Message Protocol (ICMP) echo request and reply messages, to validate an LSP. The destination IP address of the MPLS echo request packet is different from the address used to select the label stack.

The MPLS LSP Traceroute feature is used to isolate the failure point of an LSP. It is used for hop-by-hop fault localization and path tracing. The MPLS LSP Traceroute feature relies on the expiration of the Time to Live (TTL) value of the packet that carries the echo request. When the MPLS echo request message hits a transit node, it checks the TTL value and if it is expired, the packet is passed to the control plane, else the message is forwarded. If the echo message is passed to the control plane, a reply message is generated based on the contents of the request message.

The MPLS LSP Tree Trace (traceroute multipath) operation is also supported for IGP Prefix SID. MPLS LSP Tree Trace provides the means to discover all possible equal-cost multipath (ECMP) routing paths of an LSP to reach a destination Prefix SID. It uses multipath data encoded in echo request packets to query for the load-balancing information that may allow the originator to exercise each ECMP. When the packet TTL expires at the responding node, the node returns the list of downstream paths, as well as the multipath information that can lead the operator to exercise each path in the MPLS echo reply. This operation is performed repeatedly for each hop of each path with increasing TTL values until all ECMP are discovered and validated.

MPLS echo request packets carry Target FEC Stack sub-TLVs. The Target FEC sub-TLVs are used by the responder for FEC validation. The IGPIPv4 prefix sub-TLV has been added to the Target FEC Stack sub-TLV. The IGP IPv4 prefix sub-TLV contains the prefix SID, the prefix length, and the protocol (IS-IS or OSPF).

The network node which advertised the Node Segment ID is responsible for generating a FEC Stack Change sub-TLV with pop operation type for Node Segment ID, regardless of whether penultimate hop popping (PHP) is enabled or not.

The format is as below for IPv4 IGP-Prefix Segment ID:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          IPv4 Prefix                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Prefix Length  |    Protocol    |        Reserved              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The format is as below for IPv6 IGP-Prefix Segment ID:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
|                          IPv6 Prefix                          |
|                                                               |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Prefix Length  |    Protocol    |             Reserved         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

MPLS Ping and Traceroute for IGP-Adjacency Segment ID

The network node that is immediate downstream of the node which advertised the Adjacency Segment ID is responsible for generating FEC Stack Change sub-TLV for "POP" operation for Adjacency Segment ID.

The format is as below for IGP-adjacency SID:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Adj. Type   |    Protocol    |         Reserved             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Local Interface ID (4 or 16 octets)             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Remote Interface ID (4 or 16 octets)            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                                               ~
|         Advertising Node Identifier (4 or 6 octets)           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                                               ~
|            Receiving Node Identifier (4 or 6 octets)          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Configuring Segment Routing MPLS Traffic Engineering for MPLS Ping and Traceroute

```
ping mpls traffic-eng tunnel <tun-id>
[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]
[output {interface <tx-interface>} [nexthop <nexthop ip addr>]]
[{dsmap | ddmap [l2ecmp]} [hashkey {none | {ipv4 | ipv4-label-set {bitmap <bitmap_size>}}]

traceroute mpls [multipath] traffic-eng <tunnel-interface>
[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]]
[pad-tlv]]
[output {interface <tx-interface>} [nexthop <nexthop ip addr>]]
[flags {fec | ttl}]
[hashkey ipv4 | ipv4-label-set {bitmap <bitmap_size>}]
```

Configuring Segment Routing MPLS IGP for MPLS Ping and Traceroute

```
ping mpls ipv4 <prefix/prefix_length> [fec-type [ldp | bgp | generic | isis | ospf]]
[sr-path-type [ip | sid | strict-sid]]
[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]
[output {interface <tx-interface>} [nexthop <nexthop ip addr>]]
[{dsmap | ddmap [l2ecmp]} [hashkey {none | {ipv4 | ipv4-label-set {bitmap <bitmap_size>}}]

traceroute mpls [multipath] ipv4 <prefix/prefix_length> [fec-type [ldp | bgp | generic |
isis | ospf]] [sr-path-type [ip | sid | strict-sid]]
[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]]
[pad-tlv]]
[output {interface <tx-interface>} [nexthop <nexthop ip addr>]]
[flags {fec | ttl}]
[hashkey ipv4 | ipv4-label-set {bitmap <bitmap_size>}]
```

- fec-type: IPv4 Target FEC type, use head end auto detected FEC type by default.
- sr-path-type: Segment routing path type selection algorithm. Use IP imposition path, when option is specified.

Additional References for Segment Routing OAM Support

Related Documents

Related Topic	Document Title
Cisco IOS commands	Cisco IOS Master Commands List, All Releases

Feature Information for Segment Routing OAM Support

The following table provides release information about the feature or features described in this module. This table lists only the software release that introduced support for a given feature in a given software release train. Unless noted otherwise, subsequent releases of that software release train also support that feature.

Use Cisco Feature Navigator to find information about platform support and Cisco software image support. To access Cisco Feature Navigator, go to www.cisco.com/go/cfn. An account on Cisco.com is not required.

Table 12: Feature Information for Segment Routing OAM Support

Feature Name	Releases	Feature Information
Segment Routing OAM Support	Cisco IOS XE Release 3.17 S	The Segment Routing OAM feature provides support for Nil-FEC (forwarding equivalence classes) LSP Ping and Traceroute functionality. The Nil-FEC LSP ping and traceroute operation are simply extension of regular MPLS ping and trace route.

CHAPTER 14

Using Seamless BFD with Segment Routing

The Segment Routing TE feature provides information support for Seamless Bidirectional Forwarding Detection (S-BFD).

Restrictions For Using Seamless BFD with Segment Routing

Restrictions for Seamless-Birdirectional Forwarding (S-BFD)

- Seamless-Birdirectional Forwarding (S-BFD) supporting IPv4 only for segment routing traffic engineering (SR-TE). IPv6 is not supported.
- Single hop S-BFD session is only supported.
- RSVP-TE does not support S-BFD.

Information About Seamless BFD with Segment Routing

Bidirectional Forwarding Detection and Seamless-Bidirectional Forwarding Detection (S-BFD)

Bidirectional Forwarding Detection (BFD) is a detection protocol designed to provide fast forwarding path failure detection times for all media types, encapsulations, topologies, and routing protocols.

BFD provides a consistent failure detection method for network administrators, in addition to fast forwarding path failure detection. Because the network administrator can use BFD to detect forwarding path failures at a uniform rate, rather than the variable rates for different routing protocol hello mechanisms, network profiling and planning will be easier, and reconvergence time will be consistent and predictable.

Seamless Bidirectional Forwarding Detection (S-BFD), is a simplified mechanism for using BFD with a large proportion of negotiation aspects eliminated, thus providing benefits such as quick provisioning, as well as improved control and flexibility for network nodes initiating path monitoring.

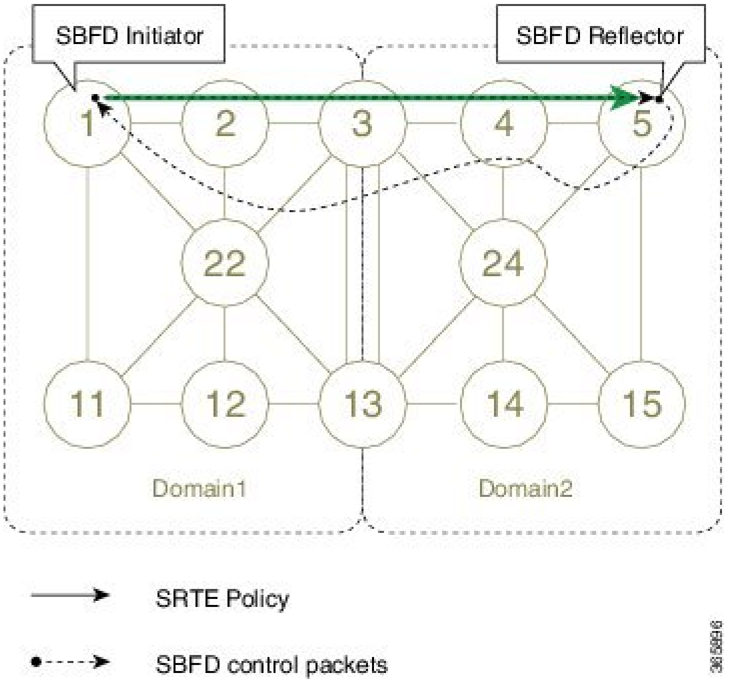
If SBFD session fails, S-BFD brings down the SR-TE session. S-BFD also provides faster session bring up due to less control packets exchange. S-BFD is associated with SR-TE to bring a session up quickly. The BFD state is only maintained at head end thereby reducing overhead.

S-BFD implements support for RFC 7880, RFC 7881 on segment routing.

Initiators and Reflectors

SBFD runs in an asymmetric behavior, using initiators and reflectors. The following figure illustrates the roles of an SBFD initiator and reflector.

Figure 19: SBFD Initiator and Reflector



The initiator is an SBFD session on a network node that performs a continuity test to a remote entity by sending SBFD packets. The initiator injects the SBFD packets into the segment-routing traffic-engineering (SRTE) policy. The initiator triggers the SBFD session and maintains the BFD state and client context.

The reflector is an SBFD session on a network node that listens for incoming SBFD control packets to local entities and generates response SBFD control packets. The reflector is stateless and only reflects the SBFD packets back to the initiator.

A node can be both an initiator and a reflector, thereby allowing you to configure different SBFD sessions.

S-BFD can be enabled and supported for SR-TE IPv4, but IPv6 is not supported. For SR-TE, S-BFD control packets are label switched in forward and reverse direction. For S-BFD, the tail end is the reflector node. Other nodes cannot be a reflector. When using S-BFD with SR-TE, if the forward and return directions are label switched paths, S-BFD need not be configured on the reflector node.

How to Configure Seamless BFD with Segment Routing

Configuring Seamless-Bidirectional Forwarding Detection (S-BFD) for Segment Routing

S-BFD must be enabled on both initiator and reflector nodes.

Note When using S-BFD with SR-TE, if the forward and return directions are label switched paths, S-BFD need not be configured on the reflector node.

Enabling Seamless Bidirectional Forwarding Detection (S-BFD) on the Reflector Node

Perform this task to configure S-BFD on the reflector node.

```
sbfd local-discriminator 55.55.55.55
```

Enabling Seamless Bidirectional Forwarding Detection (S-BFD) on the Initiator Node

Perform this task to configure S-BFD on the initiator node.

```
bfd-template single-hop ABC
 interval min-tx 300 min-rx 300 multiplier 10
```

Enabling Segment Routing Traffic Engineering Tunnel with Seamless-Bidirectional Forwarding (S-BFD)

```
interface Tunnel56
 ip unnumbered Loopback11
 tunnel mode mpls traffic-eng
 tunnel destination 55.55.55.55 */IP address of Reflector node/*
 tunnel mpls traffic-eng path-option 1 dynamic segment-routing
 tunnel mpls traffic-eng bfd sbfd ABC
!
end
```

Verifying S-BFD Configuration

SUMMARY STEPS

1. **show mpls traffic-engineering tunnel** *tunnel-name*
2. **show bfd neighbors**

DETAILED STEPS

Step 1 **show mpls traffic-engineering tunnel** *tunnel-name*

Verifies the SR TE state and the S-BFD session state.

Example:

```
Router# sh mpls traffic-eng tunnel tunnel 56

    Name: R1_t56                              (Tunnel56) Destination: 55.55.55.55
      Status:
        Admin: up         Oper: up     Path: valid       Signalling: connected
        path option 1, (SEGMENT-ROUTING) type dynamic (Basis for Setup, path weight 12)

      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
        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

      SBFD configured with template: ABC
        Session type: CURRENT        State: UP       SBFD handle: 0x3
        LSP ID: 1
        Last uptime duration: 3 minutes, 35 seconds
        Last downtime duration: --
          Active Path Option Parameters:
        State: dynamic path option 1 is active
        BandwidthOverride: disabled  LockDown: disabled  Verbatim: disabled
      Node Hop Count: 2
          History:
        Tunnel:
          Time since created: 4 minutes, 3 seconds
          Number of LSP IDs (Tun_Instances) used: 1
        Current LSP: [ID: 1]
          Uptime: 3 minutes, 36 seconds
      Tun_Instance: 1
      Segment-Routing Path Info (isis  level-2)
        Segment0[Link]: 12.12.12.1 - 12.12.12.2, Label: 48
        Segment1[Link]: 25.25.25.2 - 25.25.25.5, Label: 35 !
```

Step 2 **show bfd neighbors**

Verifies that BFD neighbors are established properly.

Example:

```
Router# show bfd neighbors

    MPLS-TE SR Sessions
    Interface      LSP ID(Type)                    LD/RD         RH/RS     State
    Tunnel56          1 (SR)                   4097/926365495  Up        Up
```

Additional References for Seamless BFD with Segment Routing

Related Documents

Related Topic	Document Title
Cisco IOS commands	Cisco IOS Master Command List, All Releases
Segment Routing Traffic Engineering configuration	*Segment Routing -Traffic Engineering*

Table 13: Standards and RFC

Standard/RFC	Title
draft-akiya-bfd-seamless-base-03	Seamless Bidirectional Forwarding Detection (S-BFD)
draft-ietf-isis-segment-routing-extensions-07	IS-IS Extensions for Segment Routing
draft-ietf-spring-segment-routing-09	Segment Routing Architecture
RFC 7880	Seamless Bidirectional Forwarding Detection (S-BFD)
RFC 7881	Seamless Bidirectional Forwarding Detection (S-BFD) for IPv4, IPv6, and MPLS

Feature Information for Seamless BFD with Segment Routing

The following table provides release information about the feature or features described in this module. This table lists only the software release that introduced support for a given feature in a given software release train. Unless noted otherwise, subsequent releases of that software release train also support that feature.

Use Cisco Feature Navigator to find information about platform support and Cisco software image support. To access Cisco Feature Navigator, go to www.cisco.com/go/cfn. An account on Cisco.com is not required.

Table 14: Feature Information for Segment Routing TE Feature

Feature Name	Releases	Feature Information
Segment Routing TE Feature	Cisco IOS XE Denali 16.4.1	Seamless Bidirectional Forwarding Detection (S-BFD), is a simplified mechanism for using BFD with a large proportion of negotiation aspects eliminated, thus providing benefits such as quick provisioning, as well as improved control and flexibility for network nodes initiating path monitoring. The following commands were introduced or modified: **address-family ipv4 strict-spf**, **bfd-template single-hop**, **index range**, **sbfd local-discriminator**, **show bfd neighbor**, **show isis segment-routing**, **show mpls forwarding-table**, **show mpls traffic tunnel**, **show mpls traffic-engineering**.

CHAPTER 15

Using SSPF with Segment Routing

The Segment Routing TE feature provides information support for the Strict Shortest Path First (SPF).

Information About SSPF with Segment Routing

Strict Shortest Path First

Segment Routing supports the following two algorithms:

- Algorithm 0: This is a Shortest Path First (SPF) algorithm based on link metric. This shortest path algorithm is computed by the Interior gateway protocol (IGP).

- Algorithm 1: This is a Strict Shortest Path First (SSPF) algorithm based on link metric. The algorithm 1 is identical to algorithm 0 but requires that all nodes along the path honor the SPF routing decision. Local policy does not alter the forwarding decision. For example, a packet is not forwarded through locally engineered path.

Different SIDs are associated with the same prefix for each algorithm.

Strict Shortest Path First is supported by default - but strict SIDs must be configured for at least one node address on each node supporting Segment Routing.

Approaches for Configure Strict Shortest Path First

The two approaches to configure Strict SFP are as follows:

- Using the **connect-prefix-sid-map** command—Strict SFP is configured globally on all the nodes. For a network to be Strict SFP-aware (that is, for ISIS to populate Strict SPF), all nodes must be configured with a local Strict SFP SID.

- Using Segment-routing Mapping Server—One node in the network is configured as mapping server and the remaining nodes act as a client.

How to Configure SSPF with Segment Routing

Configuring Strict Shortest Path First (SPF)

Enabling Strict Shortest Path First Using the connect-prefix-sid-map command

Enabling Shortest Path First on a Provider-Edge Device

When enabling Strict Shortest Path First using the **connect-prefix-sid-map** command, the Strict Shortest Path First (SPF) must be configured on the provider-edge device first and then on the node devices. The following is a sample configuration code snippet to enable Strict Shortest Path First on a provider-edge device.

```
segment-routing mpls
connected-prefix-sid-map
  address-family ipv4
   10.10.10.10/32 index 100 range 1
  exit-address-family
  address-family ipv4 strict-spf
   10.10.10.10/32 index 1000 range 1 ------------------configure strict SPF locally
  exit-address-family
```

Enabling Shortest Path First on a Node Device

The following is a sample configuration code snippet to enable Strict Shortest Path First on a node in the network and must be enabled on all nodes in a network.

```
segment-routing mpls
connected-prefix-sid-map
  address-family ipv4
   20.20.20.20/32 index 110 range 1
  exit-address-family
  address-family ipv4 strict-spf
   20.20.20.20/32 index 1100 range 1
  exit-address-family
```

Enabling Strict Shortest Path First Using Segment Routing Mapping Server

Configuring a Node as Segment Routing Mapping Server

The following is a sample configuration code snippet to configure a node as Segment Routing Mapping Server.

```
segment-routing mpls
mapping-server
  prefix-sid-map
   address-family ipv4
    10.10.10.10/32 index 100 range 1
    20.20.20.20/32 index 110 range 1
    30.30.30.30/32 index 120 range 1
    40.40.40.40/32 index 130 range 1
    50.50.50.50/32 index 140 range 1
   exit-address-family
   address-family ipv4 strict-spf
    10.10.10.10/32 index 1000 range 1
    20.20.20.20/32 index 1100 range 1
    30.30.30.30/32 index 1200 range 1
    40.40.40.40/32 index 1300 range 1
    50.50.50.50/32 index 1400 range 1
    100.100.100.100/32 index 2000 range 1
   exit-address-family
```

Configuring the Segment Routing Mapping Server to Advertise and Receive Local Prefixes

The following is a sample configuration code snippet to configure a Segment Routing Mapping Server to advertise and receive local prefixes.

```
router isis SR
segment-routing mpls
 segment-routing prefix-sid-map advertise-local
 segment-routing prefix-sid-map receive
```

Verifying ISIS Advertises the SIDs

The following is a sample configuration code snippet to verify that ISIS advertises the SIDs.

```
Router# show isis segment-routing prefix-sid-map advertise strict-spf
Tag SR:
IS-IS Level-1 advertise prefix-sid maps:
Prefix               SID Index    Range        Flags
10.10.10.10/32       1000         1
20.20.20.20/32       1100         1
30.30.30.30/32       1200         1
40.40.40.40/32       1300         1
50.50.50.50/32       1400         1
100.100.100.100/32   2000         1
Tag SR:
IS-IS Level-2 advertise prefix-sid maps:
Prefix               SID Index    Range        Flags
10.10.10.10/32       1000         1
20.20.20.20/32       1100         1
30.30.30.30/32       1200         1
40.40.40.40/32       1300         1
50.50.50.50/32       1400         1
100.100.100.100/32   2000         1
```

The following is a sample configuration code snippet to verify that a provider-edge device receives Strict Shortest Path First SID from SRMS Server.

```
Router# show isis segment-routing prefix-sid-map receive strict-spf

Tag SR:
IS-IS Level-1 receive prefix-sid maps:
Host                   Prefix               SID Index    Range        Flags
P1                     10.10.10.10/32       1000         1
                       20.20.20.20/32       1100         1
                       30.30.30.30/32       1200         1
                       40.40.40.40/32       1300         1
                       50.50.50.50/32       1400         1
                       100.100.100.100/32   2000         1
Tag SR:
IS-IS Level-2 receive prefix-sid maps:
Host                   Prefix               SID Index    Range        Flags
P1                     10.10.10.10/32       1000         1
                       20.20.20.20/32       1100         1
                       30.30.30.30/32       1200         1
                       40.40.40.40/32       1300         1
                       50.50.50.50/32       1400         1
                       100.100.100.100/32   2000         1
```

Additional References for SSPF with Segment Routing

Related Documents

Related Topic	Document Title
Cisco IOS commands	Cisco IOS Master Command List, All Releases
Segment Routing Traffic Engineering configuration	*Segment Routing -Traffic Engineering*

Feature Information for SSPF with Segment Routing

The following table provides release information about the feature or features described in this module. This table lists only the software release that introduced support for a given feature in a given software release train. Unless noted otherwise, subsequent releases of that software release train also support that feature.

Use Cisco Feature Navigator to find information about platform support and Cisco software image support. To access Cisco Feature Navigator, go to www.cisco.com/go/cfn. An account on Cisco.com is not required.

Table 15: Feature Information for Segment Routing SSPF Feature

Feature Name	Releases	Feature Information
Segment Routing TE Feature	Cisco IOS XE Denali 16.4.1	The Segment Routing TE feature provides information support for the Strict Shortest Path First (SPF).. The following commands were introduced or modified: **address-family ipv4 strict-spf**, **bfd-template single-hop**, **index range**, **sbfd local-discriminator**, **show bfd neighbor**, **show isis segment-routing**, **show mpls forwarding-table**, **show mpls traffic tunnel**, **show mpls traffic-engineering**.

CHAPTER 16

Dynamic PCC

The Stateful Path Computation Element Protocol(PCEP) enables a router to report and optionally delegate Label Switched Paths (LSPs) which is established using either Resource Reservation Protocol (RSVP) protocol or Segment Routing Traffic Engineering (SR-TE) to a stateful Path Computation Element (PCE).

An LSP delegated to a PCE can be updated by the PCE and a stateful PCE can compute and provide the path of an LSP to the Path Computation Client (PCC).

SR-TE and RSVP-TE LSPs require link-state routing protocols such as OSPF or ISIS to distribute and learn traffic engineering topology. A stateful PCE can learn the traffic engineering topology through BGP Link-State protocol. You can use the verbatim path option in the case when all or some of the intermediate nodes in a network do not support IGP extensions for TE.

Information About Dynamic PCC

Path Computation Element Protocol Functions

A Path Computation Element Protocol (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.

Redundant Path Computation Elements

For redundancy it may be required to deploy redundant PCE servers. A PCC uses precedence to select stateful PCEs for delegating LSPs. Precedence can take any value between 0 and 255. The default precedence value is 255. When there are multiple stateful PCEs with active PCEP session, PCC chooses the PCE with the lowest precedence value. In case where primary PCE server session goes down, PCC router re-delegates all tunnels to next available PCE server. You can use the following CLIs in the case of redundant PCEs:

```
R2(config)#mpls traffic-eng pcc peer 77.77.77.77 source 22.22.22.22 precedence 255
R2(config)#mpls traffic-eng pcc peer 88.88.88.88  source 22.22.22.22 precedence 100
!
end
```

In the above example PCE server with IP address 88.88.88.88 is the primary PCE server since it has lower precedence value.

How to Configure Dynamic PCC

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
```

Note **mpls traffic-eng pcc report-all** is not mandatory for PCE/PCC basic operational delegated tunnels. It is required to report locally calculated LSPs to the PCE server.

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
```

Configuring Dynamic PCC With Verbatim Path Option

To enable Dynamic PCC with verbatim path option, use the following CLI under the SR-TE tunnel interface:

```
R1#
interface Tunnel2
```

```
ip unnumbered Loopback11
tunnel mode mpls traffic-eng
tunnel destination 66.66.66.66
tunnel mpls traffic-eng autoroute destination
tunnel mpls traffic-eng path-option 1 dynamic segment-routing pce verbatim
```

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 mpls traffic-eng tunnels tunnel 1** command which shows the LSP details.

```
Device# show mpls traffic-eng tunnels 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
```

Verifying Verbatim Path Option With Dynamic PCC

To verify proper operation with verbatim path option, use the following command:

```
R1#sh mpl tr tun tun 2
Name: R1_t2                               (Tunnel2) Destination: 66.66.66.66
  Status:
    Admin: up         Oper: up     Path: valid       Signalling: connected
    path option 1, (SEGMENT-ROUTING) (PCE) type dynamic (verbatim) (Basis for Setup)

  Config Parameters:
    Bandwidth: 0        kbps (Global)  Priority: 7  7   Affinity: 0x0/0xFFFF
    Metric Type: TE (interface)
    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
    AutoRoute destination: enabled
    auto-bw: disabled
    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: enabled

  PCEP Info:
    Delegation state: Working: yes   Protect: no
    Delegation peer: 77.77.77.77
    Working Path Info:
      Request status: processed
      Created via PCRep message from PCE server: 77.77.77.77
      PCE metric: 4, type: TE
    Reported paths:
      Tunnel Name: Tunnel2_w
       LSPs:
        LSP[0]:
         source 11.11.11.11, destination 66.66.66.66, tunnel ID 2, LSP ID 1
         State: Admin up, Operation active
         Binding SID: 17
         Setup type: SR
         Bandwidth: requested 0, used 0
         LSP object:
           PLSP-ID 0x80002, flags: D:0 S:0 R:0 A:1 O:2
         ERO:
           SID[0]: Adj, Label 24, NAI: local 12.12.12.1 remote 12.12.12.2
           SID[1]: Adj, Label 26, NAI: local 25.25.25.2 remote 25.25.25.5
           SID[2]: Adj, Label 22, NAI: local 56.56.56.5 remote 56.56.56.6

  History:
    Tunnel:
      Time since created: 39 days, 19 hours, 9 minutes
      Time since path change: 1 minutes, 3 seconds
      Number of LSP IDs (Tun_Instances) used: 1
    Current LSP: [ID: 1]
      Uptime: 1 minutes, 3 seconds
  Tun_Instance: 1
  Segment-Routing Path Info (IGP information is not used)
    Segment0[Link]: 12.12.12.1 - 12.12.12.2, Label: 24
    Segment1[Link]: 25.25.25.2 - 25.25.25.5, Label: 26
    Segment2[Link]: 56.56.56.5 - 56.56.56.6, Label: 22
!
end
```

Additional References for Dynamic PCC

Related Documents

Related Topic	Document Title
Cisco IOS commands	Cisco IOS Master Commands List, All Releases

Feature Information for Dynamic PCC

The following table provides release information about the feature or features described in this module. This table lists only the software release that introduced support for a given feature in a given software release train. Unless noted otherwise, subsequent releases of that software release train also support that feature.

Use Cisco Feature Navigator to find information about platform support and Cisco software image support. To access Cisco Feature Navigator, go to www.cisco.com/go/cfn. An account on Cisco.com is not required.

Table 16: Feature Information for Dynamic PCC

Feature Name	Releases	Feature Information
Dynamic PCC	Cisco IOS XE Everest 16.6.1	The Dynamic Path Computation Client (PCC) feature supports an LSP delegated to a Path Computation Element (PCE).Dynamic PCC aupports both RSVP-TE and SR-TE. The following commands were added or modified: **show pce client peer detail**, **show mpls traffic-eng tunnels tunnel 1**, **show pce client lsp detail**, **show pce lsp detail**.