7750 Service Router
7950 Extensible Routing System

Releases up to 25.10.R3

Segment Routing and PCE Advanced Configuration Guide for MD CLI

3HE 20805 AAAF TQZZA
Edition: 01
March 2026

Nokia is committed to diversity and inclusion. We are continuously reviewing our customer documentation and consulting with standards bodies to ensure that terminology is inclusive and aligned with the industry. Our future customer documentation will be updated accordingly.

Table of contents

List of tables

List of figures

Preface

About This Guide

Each Advanced Configuration Guide is organized alphabetically and provides feature and configuration explanations, CLI descriptions, and overall solutions. The Advanced Configuration Guide chapters are written for and based on several Releases, up to 25.10.R3. The Applicability section in each chapter specifies on which release the configuration is based.

The Advanced Configuration Guides supplement the user configuration guides listed in the *7450 ESS, 7750 SR, and 7950 XRS Guide to Documentation*.

Audience

This manual is intended for network administrators who are responsible for configuring the routers. It is assumed that the network administrators have a detailed understanding of networking principles and configurations.

BGP Segment Routing Using the Prefix SID Attribute

This chapter describes BGP Segment Routing using the prefix SID attribute.

Topics in this chapter include:

Applicability

This chapter was initially based on SR OS Release 20.2.R1, but the MD-CLI configuration in the current edition is based on SR OS Release 25.10.R2. BGP Segment Routing (SR) is supported in SR OS Release 19.10.R1, and later.

Overview

Segment Routing (SR) has become a foundational technology for Software-Defined Networking (SDN) in Wide Area Networks (WANs). Also, SR is being extended beyond WAN borders into Data Centers (DCs).

SR allows an ingress node to route a packet from the source, by prepending an SR header containing an ordered list of segment identifiers (SIDs). A SID represents a topological or service-based instruction. A SID can have a local meaning for one specific node, or a global meaning within the SR domain, such as the instruction to forward a packet on the Equal-Cost Multipath (ECMP) aware shortest path to reach some prefix.

In WAN networks, infrastructure IP reachability is nearly always conveyed by an IGP protocol, such as OSPF and IS-IS, but in large-scale DCs, BGP has become the protocol of choice. In a typical DC design, BGP is used for endpoint reachability, as follows:

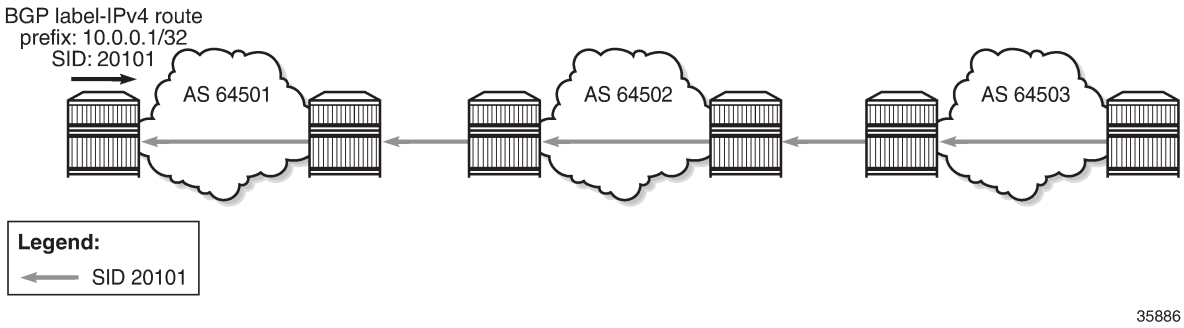
- Each node (Top of Rack (TOR), leaf, spine, and so on) has its own Autonomous System (AS).
- Each node has an EBGP session to each of its directly connected peers.
- Each node originates the IPv4 (or IPv6) address of its loopback interface into BGP and announces it to its neighbors.

To extend SR-MPLS into DCs that use this type of BGP design, the SR OS nodes must advertise their loopback IP prefix in a BGP labeled-unicast (BGP-LU) IPv4 route with a prefix SID attribute. The prefix SID attribute is ignored when attached to other types of BGP routes, including BGP-LU IPv6 routes, but it is still being propagated.

A BGP prefix SID is always a global SID within the SR domain and identifies an instruction to forward the packet along the ECMP-aware BGP-computed best paths to reach the prefix. The BGP prefix SID attribute can also help to create SR paths that transit across multiple administrative domains that do not share IGP SR topology information.

Figure 1: BGP-LU IPv4 route with prefix SID BGP path attribute shows a node in AS 64501 advertising a BGP-LU IPv4 route for prefix 10.0.0.1/32 with SID 20101. The SR-capable nodes forward packets with SID 20101 via the best BGP path to 10.0.0.1, using any of the available multipaths computed by BGP.

Figure 1: BGP-LU IPv4 route with prefix SID BGP path attribute

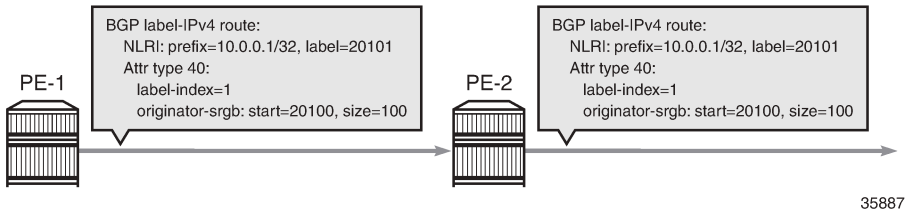


The BGP prefix SID attribute with type code 40 is an optional and transitive BGP path attribute, meaning that the attribute is expected to be propagated by routers that do not recognize the type value. When SR is deployed using an MPLS dataplane (SR-MPLS), the BGP prefix SID encodes:

- A 32-bit label-index Type-Length-Value (TLV) (mandatory TLV)
- An originator Segment Routing Global Block (SRGB) TLV containing one or more SRGB fields (optional TLV). If the SRGB field occurs multiple times in the SRGB TLV, the SRGB space of the ingress node consists of multiple ranges that are concatenated.

Figure 2: BGP signaling overview shows that node PE-1 exports a BGP-LU IPv4 route with prefix 10.0.0.1/32 and label 20101. The BGP prefix SID attribute is attribute type 40 and contains an SR label index of 1 and the originator SRGB with start label 20100 and size 100 (from 20100 to 20199). Node PE-2 imports the BGP-LU IPv4 route and exports it to the next node.

Figure 2: BGP signaling overview



To add, replace, or process a BGP prefix SID, SR must be administratively enabled in the **bgp** context. The BGP prefix SID range can be set to either **global** (that is, equal to the SRGB also used by SR-OSPF or SR-ISIS and defined in the **router "Base" mpls-labels sr-labels** context) or a subset of the SRGB defined by the **start-label** command in combination with **max-index**. All BGP prefix SID values must reside within the global SRGB or the **start-label** command fails. The **prefix-sid-range** is a mandatory requirement.

To originate BGP SR prefixes, two policies are required with an **sr-label-index** action, which may or may not be identical:

- **route-table-import policy-name <policy-name>** used to populate a local BGP-SR table with an SR label index
- **export policy [<policy>]** to advertise a prefix to a neighbor with an SR label index

In the example topology used in this chapter, the import and export policies are identical and have an **action** entry with **action-type accept** with **sr-label-index** with **value 1**, so on PE-1, the prefix SID for the prefix 10.0.0.1/32 equals 20101, which is the sum of the start label for the prefix SID range 20100 and the SR label index 1.

A unique label index value must be assigned to each different IPv4 prefix that is advertised with a BGP prefix SID. However, in case of a conflict with another SR-programmed Label Forwarding Instance Base (LFIB) entry, the conflict situation is addressed as follows:

- If the conflict is with another BGP-LU IPv4 route for a different prefix with a prefix SID attribute, all the conflicting BGP-LU IPv4 routes for both prefixes are advertised with normal BGP-LU labels from the dynamic label range, not from the dedicated SR label range.
- If the conflict is with an IGP route and the route-table-import policy action does not contain the **prefer-igp** in the **sr-label-index** command, the BGP-LU IPv4 route loses to the IGP route and is advertised with a normal BGP-LU label from the dynamic SR label range.
- If the conflict is with an IGP route and the route-table-import policy action contains the **prefer-igp** in the **sr-label-index** command, this is not considered a conflict and BGP uses the IGP-signaled label index to derive its advertised label. This stitches the BGP SR tunnel to the IGP SR tunnel.

Stitching of SR-ISIS or SR-OSPF to SR-BGP is one of the main advantages of implementing SR-BGP.

Any /32 BGP-LU IPv4 route containing a prefix SID attribute is resolvable and usable in the same way as /32 BGP-LU IPv4 routes without prefix SID attribute. The routes can be installed in the route table and tunnel table, have ECMP next hops or FRR backup next hops, and can be used as transport tunnels.

Receiving a /32 BGP-LU IPv4 route with prefix SID attribute does not create a tunnel in the SR database; it only creates a label swap entry when the route is re-advertised with a new next hop. This means that the first SID in any SID list of an SR policy should not be based on a BGP prefix SID because the data path would not be programmed correctly. However, the BGP prefix SID can be used as a non-first SID in any SR policy.

Each node capable of receiving and propagating the BGP prefix SID attribute can be configured with the **block-prefix-sid** command at the BGP global, group, or neighbor configuration levels to:

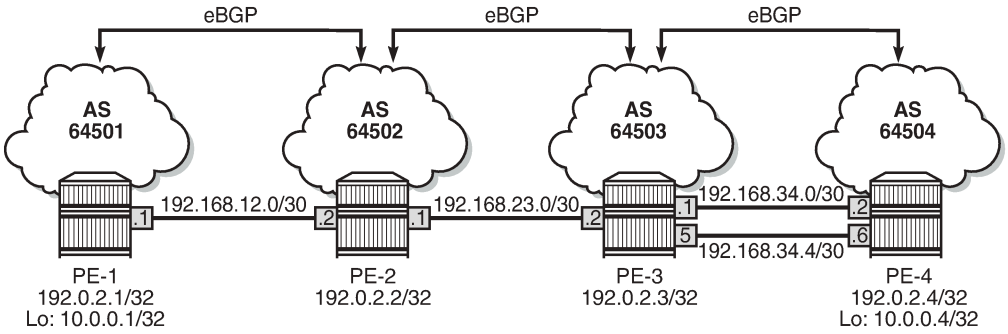
- block the propagation of the attribute outside its local SR domain
- block inbound propagation of the attribute from another SR domain

When **block-prefix-sid** applies to a BGP session, the prefix SID attribute is stripped from all sent and received routes on that session, even if the prefix SID attribute was added to the outbound routes by the local router. By default, this feature is not configured, so the prefix SID is propagated freely to and from all BGP peers.

Configuration

Figure 3: Example topology shows the example topology with four nodes in different ASs. The loopback addresses 10.0.0.1/32 on PE-1 and 10.0.0.4/32 on PE-4 are exported in BGP-LU IPv4 routes with prefix SID attribute.

Figure 3: Example topology



The initial configuration includes:

- Cards, MDAs, ports
- Router interfaces
- EBGP sessions for the label-IPv4 address family
- PE-3 and PE-4 have ECMP and **multipath max-paths** set to 2 for BGP address family **label-ipv4**

No IGP is configured, so SR-OSPF or SR-ISIS cannot be used.

Configure BGP segment routing using prefix SID

BGP SR is enabled on all PEs. Also, the SRGB is configured and the BGP SR labels are defined as a subset of the SRGB, as follows:

```
# on PE-1, PE-2, PE-3, PE-4:
configure {
    router "Base" {
        mpls-labels {
            sr-labels {
                start 20000
                end 20999
            }
        }
        bgp {
            segment-routing {
                admin-state enable
                prefix-sid-range {
                    start-label 20100
                    max-index 99
                }
            }
        }
```

It is possible to define different policies with the **sr-label-index** action for importing and exporting the prefixes, but in this example, the same policy is used. The following policy is used for exporting and importing prefix 10.0.0.1/32 on PE-1:

```
# on PE-1:
configure {
    policy-options {
        prefix-list "10.0.0.1/32" {
            prefix 10.0.0.1/32 type exact {
            }
        }
        policy-statement "prefix-sid-1" {
            entry 10 {
                from {
                    prefix-list ["10.0.0.1/32"]
                }
                action {
                    action-type accept
                    sr-label-index {
                        value 1
                    }
                }
            }
        }
    }
```

Likewise, PE-4 exports prefix 10.0.0.4/32 with SR label index value 4, resulting in a BGP prefix SID 20104 (start label 20100 + index 4 = 20104).

The **route-table-import policy-name** command is used to populate a local BGP-SR table with SR label 20101 (20100 + 1 = 20101), as follows:

```
# on PE-1:
configure {
    router "Base" {
        bgp {
            rib-management {
                label-ipv4 {
                    route-table-import {
                        policy-name "prefix-sid-1"
                    }
                }
            }
        }
```

The export policy is configured in the BGP group, as follows:

```
# on PE-1:
configure {
    router "Base" {
        bgp {
            group "EBGP" {
                family {
                    label-ipv4 true
                }
                ebgp-default-reject-policy {
                    import false
                    export false
                }
                export {
                    policy ["prefix-sid-1"]
```

```
            }
        }
        neighbor "192.168.12.2" {
            group "EBGP"
            peer-as 64502
        }
```

The following **show** commands display the BGP-SR table on the different PEs:

```
[/]
A:admin@PE-1# show router bgp sr-label

===============================================================================
BGP SR labels
Flags: B - entry has backup next-hop, E - entry has ECMP next-hops
===============================================================================
Prefix                                          Advertised  Received    Flags
                                                  Label       Label
-------------------------------------------------------------------------------
10.0.0.1/32                                     20101       -           -
10.0.0.4/32                                     20104       20104       -
-------------------------------------------------------------------------------
Total Labels allocated:   2
===============================================================================
```

```
[/]
A:admin@PE-2# show router bgp sr-label

===============================================================================
BGP SR labels
Flags: B - entry has backup next-hop, E - entry has ECMP next-hops
===============================================================================
Prefix                                          Advertised  Received    Flags
                                                  Label       Label
-------------------------------------------------------------------------------
10.0.0.1/32                                     20101       20101       -
10.0.0.4/32                                     20104       20104       -
-------------------------------------------------------------------------------
Total Labels allocated:   2
===============================================================================
```

```
[/]
A:admin@PE-3# show router bgp sr-label

===============================================================================
BGP SR labels
Flags: B - entry has backup next-hop, E - entry has ECMP next-hops
===============================================================================
Prefix                                          Advertised  Received    Flags
                                                  Label       Label
-------------------------------------------------------------------------------
10.0.0.1/32                                     20101       20101       -
10.0.0.4/32                                     20104       20104       E
-------------------------------------------------------------------------------
Total Labels allocated:   2
===============================================================================
```

```
[/]
A:admin@PE-4# show router bgp sr-label

===============================================================================
```

```
BGP SR labels
Flags: B - entry has backup next-hop, E - entry has ECMP next-hops
===============================================================================
Prefix                                          Advertised  Received    Flags
                                                  Label       Label
-------------------------------------------------------------------------------
10.0.0.1/32                                     20101       20101       E
10.0.0.4/32                                     20104       -           -
-------------------------------------------------------------------------------
Total Labels allocated:   2
===============================================================================
```

Because PE-3 and PE-4 have ECMP and BGP multipath configured, traffic flows can be sprayed over two links. The E-flag in the last column indicates that an ECMP next-hop is available for prefix 10.0.0.4/32 on PE-3 and for prefix 10.0.0.1 on PE-4.

The tunnel table on PE-1 shows that a tunnel with ID 262145 is available toward destination 10.0.0.4/32:

```
[/]
A:admin@PE-1# show router tunnel-table

===============================================================================
IPv4 Tunnel Table (Router: Base)
===============================================================================
Destination           Owner     Encap TunnelId  Pref   Nexthop        Metric
   Color
-------------------------------------------------------------------------------
10.0.0.4/32           bgp       MPLS  262145    12     192.168.12.2   1000
-------------------------------------------------------------------------------
Flags: B = BGP or MPLS backup hop available
       L = Loop-Free Alternate (LFA) hop available
       E = Inactive best-external BGP route
       k = RIB-API or Forwarding Policy backup hop
===============================================================================
```

The FP-tunnel table provides more information about the label (20104) and next hop (192.168.12.2):

```
[/]
A:admin@PE-1# show router fp-tunnel-table 1

===============================================================================
IPv4 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                 Protocol         Tunnel-ID
  Lbl/SID
    NextHop                                                  Intf/Tunnel
  Lbl/SID (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
10.0.0.4/32                                 BGP               -
  20104
    192.168.12.2                                            1/1/c1/1:100
-------------------------------------------------------------------------------
Total Entries : 1
-------------------------------------------------------------------------------
===============================================================================
```

On PE-2, two tunnels are available: one toward destination 10.0.0.1/32 with SR label 20101 and another toward destination 10.0.0.4/32 with SR label 20104:

```
[/]
A:admin@PE-2# show router fp-tunnel-table 1

===============================================================================
IPv4 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                Protocol         Tunnel-ID
  Lbl/SID
    NextHop                                               Intf/Tunnel
  Lbl/SID (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
10.0.0.1/32                                BGP               -
  20101
    192.168.12.1                                         1/1/c2/1:1000
10.0.0.4/32                                BGP               -
  20104
    192.168.23.2                                         1/1/c1/1:1000
-------------------------------------------------------------------------------
Total Entries : 2
-------------------------------------------------------------------------------
===============================================================================
```

On PE-3, three tunnels are available: one toward destination 10.0.0.1/32 with SR label 20101 and two toward destination 10.0.0.4/32 with SR label 20104.

```
[/]
A:admin@PE-3# show router fp-tunnel-table 1

===============================================================================
IPv4 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                Protocol         Tunnel-ID
  Lbl/SID
    NextHop                                               Intf/Tunnel
  Lbl/SID (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
10.0.0.1/32                                BGP               -
  20101
    192.168.23.1                                         1/1/c2/1:1000
10.0.0.4/32                                BGP               -
  20104
    192.168.34.2                                         1/1/c1/1:1000
  20104
    192.168.34.6                                         1/1/c3/1:1000
-------------------------------------------------------------------------------
Total Entries : 2
-------------------------------------------------------------------------------
===============================================================================
```

On PE-4, two tunnels are available toward destination 10.0.0.1/32 with SR label 20101:

```
[/]
A:admin@PE-4# show router fp-tunnel-table 1

===============================================================================
IPv4 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                Protocol         Tunnel-ID
  Lbl/SID
    NextHop                                               Intf/Tunnel
  Lbl/SID (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
10.0.0.1/32                                BGP               -
  20101
    192.168.34.1                                         1/1/c2/1:1000
  20101
    192.168.34.5                                         1/1/c1/1:1000
-------------------------------------------------------------------------------
Total Entries : 1
-------------------------------------------------------------------------------
===============================================================================
```

PE-1 advertised a BGP-LU IPv4 route for prefix 10.0.0.1/32 with label 20101 to PE-2. The following command on PE-2 shows the received route:

```
[/]
A:admin@PE-2# show router bgp routes 10.0.0.1/32 label-ipv4
===============================================================================
 BGP Router ID:192.0.2.2        AS:64502       Local AS:64502
===============================================================================
 Legend -
 Status codes  : u - used, s - suppressed, h - history, d - decayed, * - valid
                 l - leaked, x - stale, > - best, b - backup, p - purge,
                 w - unused-weight-only
 Origin codes  : i - IGP, e - EGP, ? - incomplete

===============================================================================
BGP LABEL-IPV4 Routes
===============================================================================
Flag  Network                                          LocalPref   MED
      Nexthop (Router)                                 Path-Id     IGP Cost
      As-Path                                                      Label
-------------------------------------------------------------------------------
u*>i  10.0.0.1/32                                      None        None
      192.168.12.1                                     None        0
      64501                                                        20101
-------------------------------------------------------------------------------
Routes : 1
===============================================================================
```

This route is advertised to PE-3 and finally to PE-4. The following command on PE-4 shows two BGP-LU IPv4 routes for prefix 10.0.0.1/32 with label 20101: one with next hop 192.168.34.1 and another one with next hop 192.168.34.5.

```
[/]
```

```
A:admin@PE-4# show router bgp routes 10.0.0.1/32 label-ipv4
===============================================================================
 BGP Router ID:192.0.2.4        AS:64504       Local AS:64504
===============================================================================
 Legend -
 Status codes  : u - used, s - suppressed, h - history, d - decayed, * - valid
                 l - leaked, x - stale, > - best, b - backup, p - purge,
                 w - unused-weight-only
 Origin codes  : i - IGP, e - EGP, ? - incomplete

===============================================================================
BGP LABEL-IPV4 Routes
===============================================================================
Flag  Network                                          LocalPref   MED
      Nexthop (Router)                                 Path-Id     IGP Cost
      As-Path                                                      Label
-------------------------------------------------------------------------------
u*>i  10.0.0.1/32                                      None        None
      192.168.34.1                                     None        0
      64503 64502 64501                                            20101
u*>i  10.0.0.1/32                                      None        None
      192.168.34.5                                     None        0
      64503 64502 64501                                            20101
-------------------------------------------------------------------------------
Routes : 2
===============================================================================
```

The detailed output for the BGP-LU IPv4 routes on PE-4 show the prefix SID attribute with index 1 and originator SRGB with start label 20100 and size 100, as follows:

```
[/]
A:admin@PE-4# show router bgp routes 10.0.0.1/32 label-ipv4 detail
===============================================================================
 BGP Router ID:192.0.2.4        AS:64504       Local AS:64504
===============================================================================
 Legend -
 Status codes  : u - used, s - suppressed, h - history, d - decayed, * - valid
                 l - leaked, x - stale, > - best, b - backup, p - purge,
                 w - unused-weight-only
 Origin codes  : i - IGP, e - EGP, ? - incomplete

===============================================================================
BGP LABEL-IPV4 Routes
===============================================================================
Original Attributes

Network        : 10.0.0.1/32
Nexthop        : 192.168.34.1
Path Id        : None
From           : 192.168.34.1
Res. Nexthop   : 192.168.34.1
Local Pref.    : n/a                   Interface Name : int-PE-4-PE-3
Aggregator AS  : None                  Aggregator     : None
Atomic Aggr.   : Not Atomic            MED            : None
AIGP Metric    : None                  IGP Cost       : 0
Connector      : None
Community      : No Community Members
Cluster        : No Cluster Members
Originator Id  : None                  Peer Router Id : 192.0.2.3
Fwd Class      : None                  Priority       : None
IPv4 Label     : 20101
Origin         : IGP
Flags          : Used Valid Best In-TTM In-RTM
```

```
Route Source   : External
AS-Path        : 64503 64502 64501
Route Tag      : 0
Neighbor-AS    : 64503
DB Orig Val    : NotFound              Final Orig Val : N/A
Source Class   : 0                     Dest Class     : 0
Add Paths Send : Default
RIB Priority   : Normal
Last Modified  : 00h04m53s
Prefix SID     : index 1, originator-srgb [20100/100]

---snip---
```

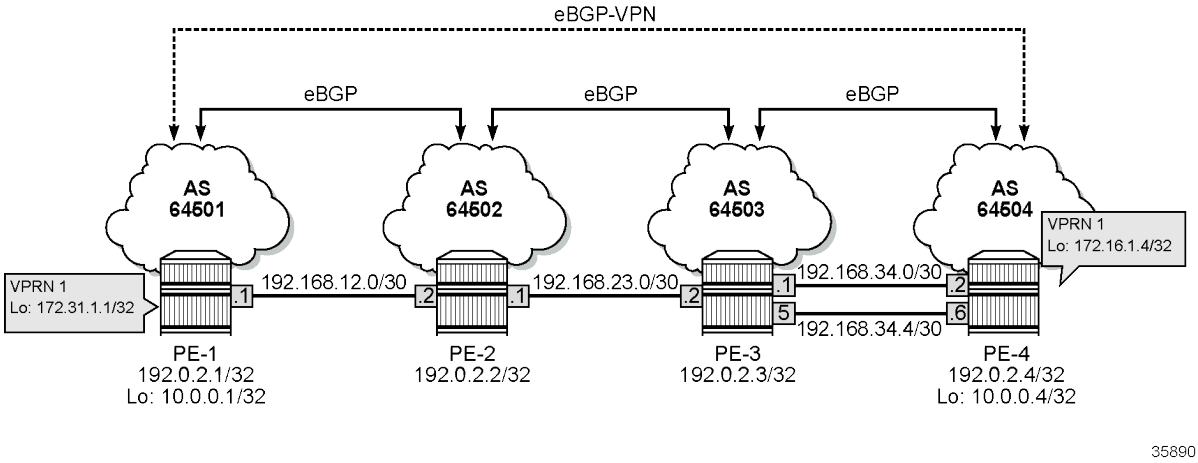
The following debug message on PE-4 shows how the prefix SID attribute is advertised in a BGP update:

```
5 2026/01/07 08:24:43.963 UTC MINOR: DEBUG #2001 Base Peer 1: 192.168.34.1
"Peer 1: 192.168.34.1: UPDATE
Peer 1: 192.168.34.1 - Received BGP UPDATE:
    Withdrawn Length = 0
    Total Path Attr Length = 66
    Flag: 0x90 Type: 14 Len: 17 Multiprotocol Reachable NLRI:
        Address Family LBL-IPV4
        NextHop len 4 NextHop 192.168.34.1
        10.0.0.1/32 Label 20101
    Flag: 0x40 Type: 1 Len: 1 Origin: 0
    Flag: 0x40 Type: 2 Len: 14 AS Path:
        Type: 2 Len: 3 < 64503 64502 64501 >
    Flag: 0xc0 Type: 40 Len: 21 Prefix-SID-attr:
       Label Index TLV Type: 1 len: 7:-
          reserved: 0 flags: 0x0  label Index: 1
       Originator SRGB TLV type: 3 len: 8 flags: 0x0:-
          start_label: 20100  num_label: 100
"
```

Configure VPRN

Figure 4: Example topology with VPRN 1 shows the example topology with a basic VPRN service to demonstrate the end-to-end control plane signaling and data plane verification.

Figure 4: Example topology with VPRN 1



A BGP multi-hop session for address family VPN-IPv4 is configured between the GRT loopback addresses 10.0.0.1/32 on PE-1 and 10.0.0.4/32 on PE-4. On PE-1, the additional BGP configuration is as follows:

```
# on PE-1:
configure {
    router "Base" {
        bgp {
            group "EBGP-VPN" {
                family {
                    vpn-ipv4 true
                }
                ebgp-default-reject-policy {
                    import false
                    export false
                }
            }
            neighbor "10.0.0.4" {
                group "EBGP-VPN"
                multihop 64
                local-address 10.0.0.1
                peer-as 64504
            }
```

In addition, the VPRN 1 service has loopback addresses 172.31.1.1/32 on PE-1 and 172.16.1.4/32 on PE-4. The configuration on PE-1 is as follows:

```
# on PE-1:
configure {
    service {
        vprn "VPRN 1" {
            admin-state enable
            service-id 1
            customer "1"
            bgp-ipvpn {
                mpls {
                    admin-state enable
                    route-distinguisher "64501:1"    # on PE-4: route-distinguisher 64504:1
                    vrf-target {
                        community "target:1:1"
                    }
                    auto-bind-tunnel {
                        resolution any
                    }
                }
            }
            interface "lo1" {
                loopback true
                ipv4 {
                    primary {
                        address 172.31.1.1            # on PE-4: address 172.16.1.4
                        prefix-length 32
                    }
                }
            }
        }
```

The configuration on PE-4 is similar.

The following VPN-IPv4 route is received on PE-1:

```
[/]
A:admin@PE-1# show router bgp routes vpn-ipv4
```

```
===============================================================================
 BGP Router ID:192.0.2.1        AS:64501       Local AS:64501
===============================================================================
 Legend -
 Status codes  : u - used, s - suppressed, h - history, d - decayed, * - valid
                 l - leaked, x - stale, > - best, b - backup, p - purge,
                 w - unused-weight-only
 Origin codes  : i - IGP, e - EGP, ? - incomplete

===============================================================================
BGP VPN-IPv4 Routes
===============================================================================
Flag  Network                                          LocalPref   MED
      Nexthop (Router)                                 Path-Id     IGP Cost
      As-Path                                                      Label
-------------------------------------------------------------------------------
u*>i  64504:1:172.16.1.4/32                            None        None
      10.0.0.4                                         None        0
      64504                                                        524286
-------------------------------------------------------------------------------
Routes : 1
===============================================================================
```

The route table for VPRN 1 on PE-1 is as follows:

```
[/]
A:admin@PE-1# show router service-name "VPRN 1" route-table

===============================================================================
Route Table (Service: 1)
===============================================================================
Dest Prefix[Flags]                             Type    Proto     Age        Pref
      Next Hop[Interface Name]                                     Metric
-------------------------------------------------------------------------------
172.16.1.4/32                                  Remote  BGP VPN   00h01m02s  170
       10.0.0.4 (tunneled:BGP)                                        1000
172.31.1.1/32                                  Local   Local     00h01m12s  0
       lo1                                                            0
-------------------------------------------------------------------------------
No. of Routes: 2
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
===============================================================================
```

Conclusion

With BGP SR, it is possible to use SR without the use of an IGP protocol (for example, to cross AS boundaries). It is also possible to stitch SR-IGP and SR-BGP tunnels together. BGP SR uses the prefix SID attribute.

BGP Signaled Segment Routing Policy

This chapter describes BGP Signaled Segment Routing Policy.

Topics in this chapter include:

Applicability

This chapter was initially based on SR OS Release 20.10.R1, but the MD-CLI configuration in the current edition is based on SR OS Release 25.10.R2.

Overview

Segment Routing (SR) allows a head-end node to steer a packet flow along a source-routed path. SR policy is a generic framework that describes the procedures and processes that a head-end node carries out when instantiating such a path. The SR policy consists of an ordered list of segments on a node, sufficient to implement a traffic-engineered path. The segments can have any type of Segment Identifier (SID), including Adjacency-SIDs, Node-SIDs, Anycast-SIDs, or Binding SIDs. The head-end can then steer traffic, using the SR policy as appropriate.

An SR policy can define one or multiple candidate paths. When explicit candidate paths are used, each path contains one or more segment lists, where each segment list contains the ordered set of segments (identified by their unique SID) required to provide the source-routed path from head-end to destination. When a candidate path contains multiple segment lists, each is assigned a weight for the purpose of weighted load-balancing. Candidate paths can be instantiated using a variety of ways, including Path Computation Element Protocol (PCEP), BGP, or local configuration. This chapter describes the use of BGP to advertise SR policy candidate paths. The term "BGP SR policy" is interchangeably used with "BGP SR TE policy".

SR policy overview

An SR policy is identified through the tuple {head-end, color, endpoint}.

- The head-end is the node where the SR policy is instantiated, and the node that is responsible for steering traffic, using the SR policy with the relevant SID stack. From the perspective of the head-end, the SR policy can be identified using the {color, endpoint} tuple.
- The color is a fundamental part of the SR policy and forms part of the Network Layer Reachability Information (NLRI). The color is a 32-bit numerical value that a head-end uses to associate the SR policy with a characteristic, such as low-latency or high-throughput.

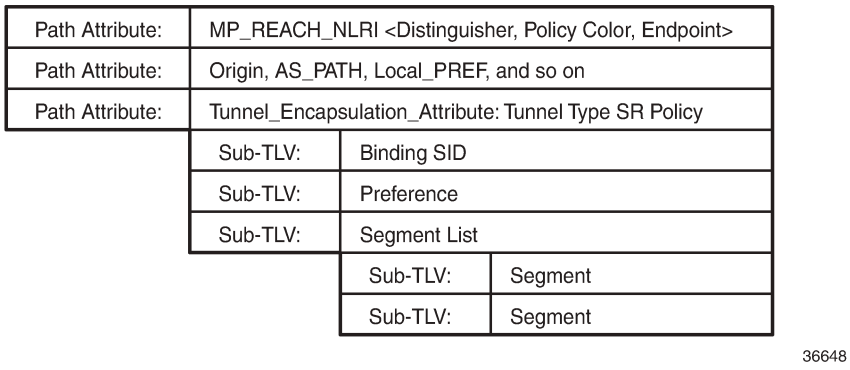
- The endpoint is the destination in the SR policy specified as an IPv4 or IPv6 address, although "wildcard" destinations can be used and are described later in this chapter.

Color is also a 32-bit transitive extended community originally defined in *draft-ietf-idr-tunnel-encaps* that can be attached to a BGP update message, in order to associate it with a corresponding SR policy. For example, if head-end H learns a BGP route R with {next-hop N, color extended community C, and VPN label V} and head-end H has a valid SR policy P to {endpoint N, color C}, it can associate BGP route R with the SR policy P. When H receives packets with a destination matching BGP route R, it forwards them using the instructions contained within SR policy P.

SR policy NLRI

The BGP address family "SR TE policy" (SAFI 73) is defined to advertise a candidate path for an SR policy in BGP and is carried in an update message using BGP multiprotocol extensions. The AFI must be IPv4 (AFI=1) or IPv6 (AFI=2). An SR policy candidate path may be advertised from a centralized controller, or it may be advertised by a router; for example, an egress router advertising paths to itself. Figure 5: SR TE policy NLRI shows the structure of the SR TE policy NLRI.

Figure 5: SR TE policy NLRI



The SR TE policy NLRI is used to identify an SR policy candidate path and, because it uses MP-BGP, it is carried in an MP_REACH/UNREACH_NLRI path attribute. The NLRI contains the color and endpoint values described previously, and a distinguisher. The distinguisher is an integer value in the range 1 to 4294967295 that serves to make the SR policy unique from an NLRI perspective. The SR TE policy NLRI uses standard BGP propagation and best-path selection; a unique distinguisher ensures that best-path selection does not unnecessarily suppress SR policy advertisements.

Multiple candidate paths can exist for an SR policy, although only one path can be selected as the best path of the SR policy and become the active path. If several candidate paths of the same SR policy (endpoint, color) are advertised via BGP SR TE policy to the same head-end, unique distinguishers for each NLRI are recommended.

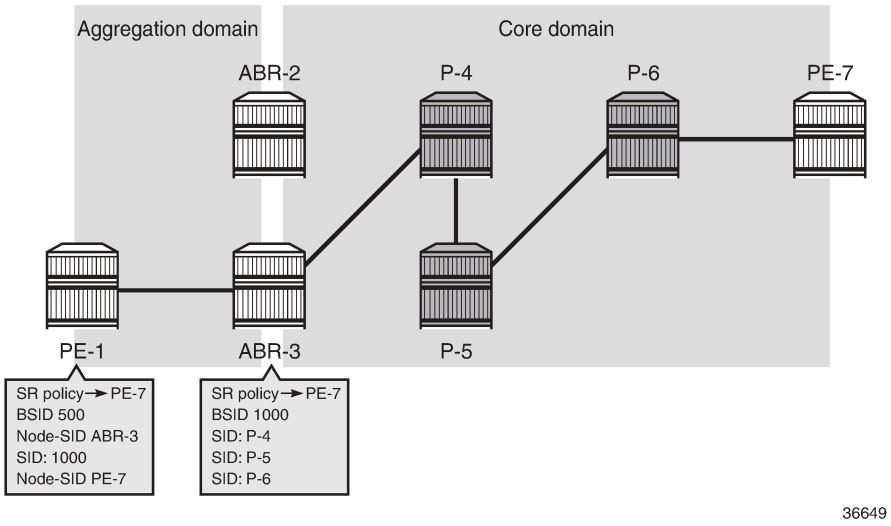
The other parameters of the SR policy candidate path are carried as sub-TLVs of the Tunnel Encapsulation Attribute (*draft-ietf-idr-tunnel-encaps*) using a tunnel-type known as "SR policy", and are described following.

Binding SID

The SR architecture defines the use of a Binding SID (BSID). A BSID is bound to an SR policy, and packets arriving at a node with an active label equal to the BSID are steered using that SR policy. This action may mean swapping the incoming active label with one or more outgoing labels representing the SR policy path.

When used in this manner, the Binding SID serves as an anchor point, sometimes referred to as a "BSID anchor", that allows one domain to be isolated from another domain. This is shown in Figure 6: Binding SID (BSID) anchor, where ABR-3 is acting as a BSID anchor between the aggregation domain and the core domain. ABR-3 has an SR policy to PE-7 with the path P-4-P-5-P-6 and with a BSID of 1000. The PE-1 resulting SR policy to PE-7 consists of the path {Node-SID ABR-3, 1000, Node-SID PE-7} and a BSID of 500. When a packet is forwarded by the SR policy on PE-1 and arrives at ABR-3, it pops the Node-SID ABR-3 label, and swaps label 1000 for the label stack {P-4, P-5, P-6} of the SR policy on ABR-3.

Figure 6: Binding SID (BSID) anchor



The BSID serves as an anchor point, which allows one domain to be isolated from the churn of another domain. If something changes in the path P-4-P-5-P-6, ABR-3 can repair the path locally without needing to change the BSID value known at PE-1. PE-1 is therefore protected from the churn in the core domain. The BSID also serves to reduce the number of segments/labels that the head-end needs to impose an end-to-end traffic-engineered path.

Segment list

A segment list sub-TLV encodes a single path toward the endpoint. Multiple segment list sub-TLVs may be included in each SR policy. Each segment list sub-TLV may contain multiple segment sub-TLVs and may carry a weight sub-TLV. Each segment sub-TLV describes a single segment in a segment list, and multiple segments may be concatenated to constitute an end-to-end path of the SR policy.

There are several types of the segment sub-TLV, allowing for the segment to be expressed as a variant of IPv4/IPv6 node address or local/remote address, and with a SID in the form of an MPLS label or IPv6 address. This chapter focuses only on the Type A encoding, which is represented as a SID in the form of

an MPLS label. The SID contained within each segment sub-TLV can be any type of SID, including Node-SID, Adjacency-SID, Anycast-SID, or Binding SID.

The optional weight sub-TLV is used to implement (weighted) load-balancing in the presence of multiple segment lists. By default, SR OS assigns a weight value of 1 to each segment list.

Preference

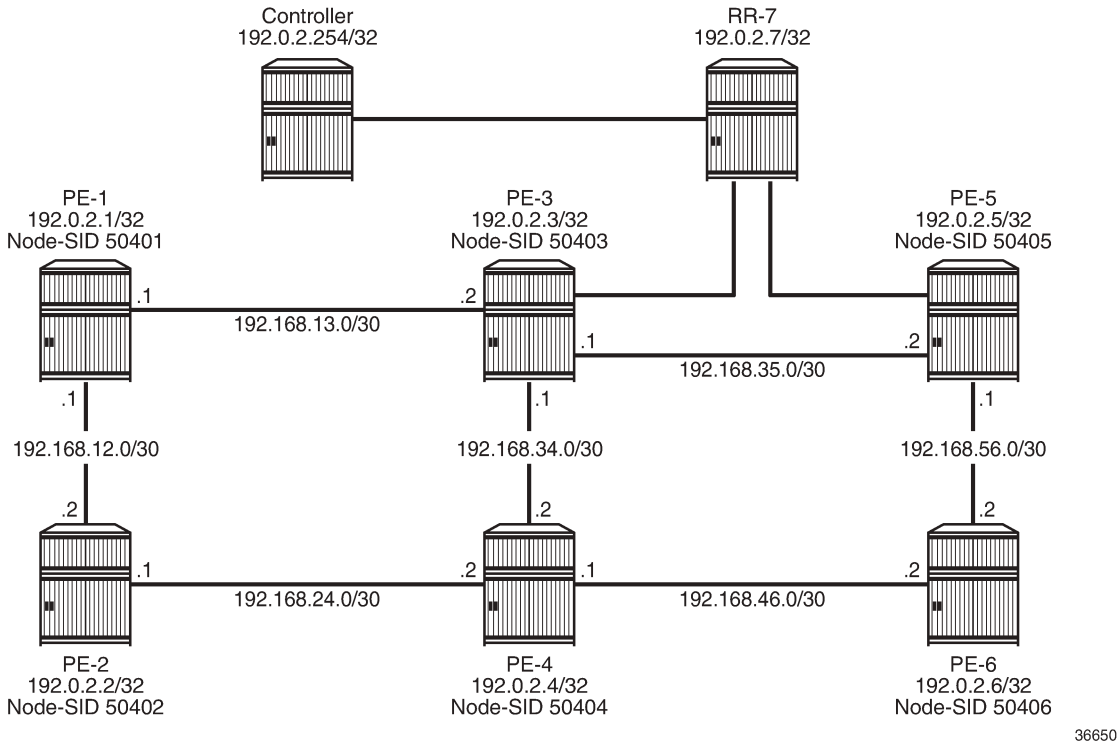
The preference sub-TLV is used to indicate the preference of a candidate path in relation to other candidate paths. Multiple candidate paths can exist in an SR policy, but only one candidate path can be selected as the best and active path. When multiple candidate paths exist that are considered valid, the candidate path with the highest preference is selected. The default value of the preference is 100. If multiple paths have the same preference, the protocol origin (PCEP, BGP, local configuration) may be considered, followed by the lower value of originator, followed by the higher value of discriminator.

Example topology

The topology in Figure 7: Example topology shows the use of BGP SR TE policy within this chapter. All PE routers within the example topology and the Route Reflector (RR-7) form part of Autonomous System 64496 and belong to the same IS-IS Level-2 area. All IGP link metrics are 100 and are symmetric. SR is enabled within the domain, and the associated Node-SIDs are shown in Figure 7: Example topology (Adj-SIDs are not shown for the purpose of clarity). The SRGB in use is {50000-54999}. All PE routers are clients of the Route Reflector for multiple address families including SR TE policy.

The example topology also has an additional router simulating a controller, which uses static routing for IP connectivity. This is the point from which SR policies are advertised into BGP, although as previously described, SR policies can be advertised into BGP by a controller or a router. The controller peers in the SR TE policy address family with the Route Reflector, which in turn reflects those routes to its clients.

Figure 7: Example topology



Configuration

An SR policy can be statically (CLI) configured locally on a head-end or dynamically learned by a head-end through BGP SR TE policy route. For SR OS to obtain an SR TE policy route, that route needs to be configured locally as a static SR policy. This chapter provides an example of the instantiation of an SR policy using static configuration on the head-end, but thereafter focuses on the instantiation of SR policies learned by it through BGP SR TE policy. The same static SR policy configuration is used regardless of whether it is for advertising that SR policy into BGP to the head-end, or applying it at that local head-end to forward traffic.

Segment Routing Local Block

A BSID may be either a local SID or a global SID. In general, and for the use-cases in this chapter, BSIDs are local SIDs, so a BSID needs to be within the range of a locally-configured Segment Routing Local Block (SRLB). SRLBs are reserved label blocks used for specific local purposes, such as SR policy BSIDs, Adjacency Set SIDs, and static Adjacency SIDs. A dedicated SRLB is required per application and has only local significance, so the same values can be used on all SR routers in the domain. Ranges for each SRLB are taken from the dynamic label range. The following configuration allocates labels 100000 to109999 to the SRLB "SRLB-BSID":

```
# on all nodes:
```

```
configure {
    router "Base" {
        mpls-labels {
            reserved-label-block "SRLB-BSID" {
                start-label 100000
                end-label 109999
            }
        }
    }
}
```

After the SRLB is defined, it is dedicated to the specific application, which in this case is SR policies. When the **reserved-label-block** is assigned, **sr-policies** must be enabled (**admin-state enable**), as follows:

```
# on all nodes:
configure {
    router "Base" {
        segment-routing {
            sr-policies {
                admin-state enable
                reserved-label-block "SRLB-BSID"
            }
        }
    }
}
```

The preceding configuration is applied to all SR routers in the domain.

Static SR policy

As previously described, SR policies can be statically (CLI) configured locally on a head-end or dynamically learned by a head-end through BGP SR TE policy route. In this section, the necessary steps are shown for the instantiation of an SR policy using static configuration locally on PE-1 as the head-end.

The following output shows the configuration of a static SR policy at PE-1 (192.0.2.1) with an endpoint of PE-5 (192.0.2.5).

```
# on PE-1:
configure {
    router "Base" {
        segment-routing {
            sr-policies {
                static-policy "PE-1-PE-5-color600" {
                    admin-state enable          # enable static SR policy
                    color 600
                    endpoint 192.0.2.5
                    head-end local
                    binding-sid 100002
                    distinguisher 600001005
                    segment-list 1 {
                        admin-state enable      # enable segment list
                        segment 1 {
                            mpls-label 50402    # node-SID PE-2
                        }
                        segment 2 {
                            mpls-label 150024   # adj-SID int-PE-2-PE-4
                        }
                        segment 3 {
                            mpls-label 150046   # adj-SID int-PE-4-PE-6
                        }
```

```
                        segment 4 {
                            mpls-label 50405    # node-SID PE-5
                        }
                    }
                }
            }
        }
    }
}
```

The static SR policy is initially created within the sr-policies context and begins by assigning a **binding-sid** of 100002. In this example, the SR policy is local to PE-1, and the BSID value is therefore within the range of the PE-1 SRLB. If this static SR policy were to be advertised into BGP, the advertised BSID value must be in the range of the SRLB configured on the target head-end.

The next three parameters are the color, distinguisher, and endpoint that constitute the SR policy NLRI. The SR policy **color** is 600, and is a 4-octet value that can be configured in the range 1 to 4294967295. The **distinguisher** is also a 4-octet value with the same range and is configured as 600001005 (representing the color plus the last octet of the head-end and endpoint addresses). As previously described, the purpose of the distinguisher is to make the SR policy unique from an NLRI perspective, such that if multiple candidate paths of the same SR policy (endpoint, color) are advertised, they are not suppressed by any BGP best-path selection algorithm.

The **endpoint** is the IPv4 or IPv6 address of the destination for the SR policy and is configured as the PE-5 address 192.0.2.5. There are special circumstances where the value 0.0.0.0 or 0::0 is allowed as an endpoint. This is referred to as color-only steering and is described later in this chapter.

The **head-end** is the target node where the SR policy is to be instantiated. If the SR policy is statically configured on the head-end for forwarding of traffic locally using that SR policy, the value **local** is used, as shown in this example. If the SR policy is configured somewhere other than on the head-end, and advertised into BGP toward the head-end, the value of the head-end parameter is the IPv4 address of that head-end. When the SR policy is advertised into BGP, the head-end address is also encoded as an IPv4 address-specific Route-Target Extended Community, which allows for potential constraining of route propagation.

The final parameter is the segment list. The preceding configuration output shows the segment list consisting of four segments, which represent the path using the following SIDs:

- Segment 1 SID is 50402, which is the Node-SID of PE-2
- Segment 2 SID is 150024, representing the PE-2 Adj-SID for the link PE-2-PE-4
- Segment 3 SID is 150046, representing the PE-4 Adj-SID for the link PE-4-PE-6
- Segment 4 SID is 50405, which is the Node-SID of PE-5.

A more optimal SID stack is achievable in this topology, but the configured segment list shows the use of both Node- and Adj-SIDs on a loose or strict hop basis. The segment list has an optional weight parameter used for load-balancing across multiple segment lists. In this example, only a single segment list exists, so the default weight value of 1 is retained.

Finally, both the segment list and the static SR policy are enabled (**admin-state enable**). The following output shows the operational state of the static SR policy. The **active** field shows whether this candidate path is the selected path in the presence of multiple candidate paths. The SR policy segment list is considered valid if the head-end is able to perform path resolution for the first SID in the segment list into one or more outgoing interfaces and next-hops. The segment 1 label is 50402, and the state is shown as **resolved-up**, indicating that this is a valid segment list.

```
[/]
```

```
A:admin@PE-1# show router segment-routing sr-policies static

===============================================================================
SR-Policies Path
===============================================================================
-------------------------------------------------------------------------------
Type            : srMpls
Active          : Yes                   Owner           : static
Operational     : Yes
Color           : 600
Head            : 0.0.0.0               Endpoint Addr   : 192.0.2.5
RD              : 600001005             Preference      : 100
BSID            : 100002
TunnelId        : 917506                Age             : 10
Origin ASN      : 0                     Origin          : 0.0.0.0
NumReEval       : 0                     LastReEvalReason: none
NumActPathChange: 0                     Last Change     : 01/29/2026 21:59:20
Maintenance Plcy:
Ret Path BFD Lbl:

Path Segment Lists:
Segment-List    : 1                     Weight          : 1
Num Segments    : 4                     Last Change     : 01/29/2026 21:45:49
  1 MPLS Label  : 50402                 State           : resolved-up
  2 MPLS Label  : 150024                State           : N/A
  3 MPLS Label  : 150046                State           : N/A
  4 MPLS Label  : 50405                 State           : N/A

===============================================================================
```

If the SR policy is considered valid, it is populated in the tunnel table with an owner of sr-policy. The entry indicates the destination and color, and always has a metric value of 0 regardless of how the SR policy is instantiated. The metric value of 0 is used because there is no effective way for the head-end to determine a more reflective value for an SR policy when learned through BGP SR TE policy or statically configured.

```
[/]
A:admin@PE-1# show router tunnel-table 192.0.2.5 protocol sr-policy

===============================================================================
IPv4 Tunnel Table (Router: Base)
===============================================================================
Destination           Owner     Encap TunnelId  Pref   Nexthop        Metric
   Color
-------------------------------------------------------------------------------
192.0.2.5/32          sr-policy MPLS  917506    14     192.0.2.2      0
   600
-------------------------------------------------------------------------------
Flags: B = BGP or MPLS backup hop available
       L = Loop-Free Alternate (LFA) hop available
       E = Inactive best-external BGP route
       k = RIB-API or Forwarding Policy backup hop
===============================================================================
```

Traffic steering using SR policies

A head-end can potentially steer traffic using an SR policy as a midpoint (or BSID anchor) or as an ingress router using color-based traffic steering:

- At a midpoint or BSID anchor, if an incoming packet has an active label that matches the BSID of a valid SR policy, the incoming label is swapped for the labels contained in the active path of that SR policy, and traffic is forwarded along that path.
- At an ingress router, if a BGP or service route is received containing a Color Extended Community with a value corresponding to a valid local SR policy, and the endpoint of that SR policy matches the next-hop of the BGP/service route, traffic is forwarded into the associated SR policy.

This sub-section discusses the use of the Color Extended Community to implement traffic steering at an ingress router, and begins with an overview of the structure of the Color Extended Community.

The Color Extended Community has two flags, known as the Color-Only (CO) bits, that allow for a head-end to optionally steer traffic using an SR policy, without the need to explicitly define an SR policy endpoint that matches the next-hop of a BGP or service route. In this case, the endpoint can be the null address (0.0.0.0 for IPv4 and 0::0 for IPv6) and traffic is steered by an SR policy based on correlation of color. Table 1: Use of CO bits describes the destination steering options based on the setting of the Color-Only (CO) bits.

Table 1: Use of CO bits

CO bits=00	CO bits=01	CO bits=10
If there is a valid SR policy (N, C), where N is the IPv4 or IPv6 endpoint address and C is a color, steer using SR policy (N, C);	If there is a valid SR policy (N, C), where N is the IPv4 or IPv6 endpoint address and C is a color, steer using SR policy (N, C);	If there is a valid SR policy (N, C), where N is the IPv4 or IPv6 endpoint address and C is a color, steer using SR policy (N, C);
Else, steer on the IGP path to the next-hop N	Else, if there is a valid SR policy (null endpoint, C) of the same address family as N, steer using SR policy (null endpoint, C);	Else, if there is a valid SR policy (null endpoint, C) of the same address family as N, steer using SR policy (null endpoint, C);
	Else, if there is any valid SR policy (any address family null endpoint, C), steer using SR policy (any null endpoint, C);	Else, if there is any valid SR policy (any address family null endpoint, C), steer using SR policy (any null endpoint, C);
	Else, steer on the IGP path to the next-hop N	Else, if there is any valid SR policy (any endpoint, C) of the same address family as N, steer using SR policy (any endpoint, C);
		Else, if there is any valid SR policy (any address family endpoint, C), steer using SR policy (any address family endpoint, C);
		Else, steer on the IGP path to the next-hop N

Per-destination traffic steering

When incoming packets match a BGP/service route with a next-hop that resolves to an SR policy, it is referred to as per-destination traffic steering. The previously configured static SR policy at PE-1 with color 600 is used to show how it is applied. A VPRN service (600) is extended between PE-1 and PE-5 with route import/export target 64496:600, and with the auto-bind-tunnel resolution-filter at PE-1 set to SR policy (the complete VPRN service configuration is not shown for conciseness).

```
# on PE-1:
configure {
    service {
        vprn "600" {
            admin-state enable
            service-id 600
            customer "1"
            autonomous-system 64496
            bgp-ipvpn {
                mpls {
                    admin-state enable
                    route-distinguisher "64496:600"
                    vrf-import {
                        policy ["vrf600-import"]
                    }
                    vrf-export {
                        policy ["vrf600-export"]
                    }
                    auto-bind-tunnel {
                        resolution filter
                        resolution-filter {
                            sr-policy true
                        }
                    }
                }
            }
            bgp {
                ---snip---
            }
            ---snip---
        }
    }
}
```

A CE router is locally connected to PE-5 and advertises prefix 10.148.5.0/24 to IPv4 BGP, which PE-5 subsequently advertises as a VPN-IPv4 route. In addition to attaching the Route-Target Extended Community to the VPN-IPv4 route, PE-5 also attaches a Color Extended Community with value 600. The PE-5 VRF export policy is shown following. When configuring the Color Extended Community, the syntax "color:co:value" is used. Therefore, in the example configuration, the CO bits are 00 and the color value is 600.

```
# on PE-5:
configure {
    policy-options {
        community "vrf600-export" {
            member "target:64496:600" { }
        }
        community "vrf600-sr-policy" {
            member "color:00:600" { }
        }
        policy-statement "vrf600-export" {
```

```
            entry 10 {
                from {
                    protocol {
                        name [bgp]
                    }
                }
                to {
                    protocol {
                        name [bgp-vpn]
                    }
                }
                action {
                    action-type accept
                    community {
                        add ["vrf600-export" "vrf600-sr-policy"]
                    }
                }
            }
        }
    }
}
```

At PE-1, the VPN-IPv4 route with {next-hop PE-5, color extended community 600} resolves to the PE-1 static SR policy with {endpoint PE-5, color 600}. It is imported into the VPRN route-table with an indication that it is resolved to the SR policy with tunnel ID 917506, which is the tunnel ID of the previously configured static SR policy. Traffic from PE-1 to PE-5 is therefore forwarded into the SR policy using the label stack defined in the segment list.

```
[/]
A:admin@PE-1# show router 600 route-table 10.148.5.0

===============================================================================
Route Table (Service: 600)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
10.148.5.0/24                                 Remote  BGP VPN   00h01m13s  170
       192.0.2.5 (tunneled:SR-Policy:917506)                        0
-------------------------------------------------------------------------------
No. of Routes: 1
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
===============================================================================
```

Color-Only traffic steering

SR OS provides support for Color-Only traffic steering using a null endpoint SR policy, but its use is limited to unlabeled BGP address families because of the following. When an egress router advertises a downstream label in a labeled BGP update (VPN-IPv4/IPv6, EVPN, BGP Labeled Unicast, and so on) that egress router needs to see that label in received packets to be able to demultiplex into the relevant service/next-hop and forward traffic toward the destination. If a head-end is forwarding traffic using an SR policy with a null endpoint, that head-end is unaware of the egress router, so cannot impose the relevant downstream-advertised BGP/service label into the label stack.

The following example shows the configuration of Color-Only traffic steering. PE-5 advertises an IPv4 prefix 172.16.5.1/32 to PE-1 with the Color Extended Community color:01:600. PE-1 intends to use the previously configured static SR policy to resolve this route. As described in Table 1: Use of CO bits, with the CO-bits set to 01, the head-end uses an SR policy with (null endpoint, C) if no valid (N, C) SR policy exists.

```
[/]
A:admin@PE-5# show router bgp routes 172.16.5.1/32 hunt | match 'Network|Nexthop|Community'
Network        : 172.16.5.1/32
Nexthop        : 192.0.2.5
Res. Nexthop   : n/a
Community      : color:01:600
```

At PE-1, the static SR policy to PE-5 is reconfigured such that the endpoint is no longer an explicit endpoint of 192.0.2.5 (PE-5), but instead uses a null endpoint (0.0.0.0).

```
# on PE-1:
configure {
    router "Base" {
        segment-routing {
            sr-policies {
                static-policy "PE-1-PE-5-color600" {
                    endpoint 0.0.0.0
                }
            }
        }
    }
}
```

Since PE-5 advertised an IPv4 BGP prefix, PE-1 also enables the use of BGP shortcuts, with a resolution filter that only permits the use of SR policy.

```
# on PE-1:
configure {
    router "Base" {
        bgp {
            next-hop-resolution {
                shortcut-tunnel {
                    family ipv4 {
                        resolution filter
                        resolution-filter {
                            sr-policy true
                        }
                    }
                }
            }
        }
    }
}
```

The tunnel table of PE-1 shows that there is a single SR policy active with a destination of 0.0.0.0/32 (null) and color 600.

```
[/]
A:admin@PE-1# show router tunnel-table protocol sr-policy

===============================================================================
IPv4 Tunnel Table (Router: Base)
===============================================================================
```

```
Destination           Owner     Encap TunnelId  Pref   Nexthop        Metric
   Color
-------------------------------------------------------------------------------
0.0.0.0/32            sr-policy MPLS  917507    14     192.0.2.2      0
   600
-------------------------------------------------------------------------------
---snip---
===============================================================================
```

The status of the IPv4 prefix 172.16.5.1/32 received from PE-5 is shown in the following output at PE-1. The output shows that the route is Used/Valid/Best, and that the resolving protocol is SR-POLICY, and the resolving NextHop is 0.0.0.0. Therefore, a BGP next-hop has been resolved to a null endpoint SR policy using the CO-bits.

```
[/]
A:admin@PE-1# show router bgp routes 172.16.5.1 detail
===============================================================================
 BGP Router ID:192.0.2.1        AS:64496       Local AS:64496
===============================================================================
 Legend -
 Status codes  : u - used, s - suppressed, h - history, d - decayed, * - valid
                 l - leaked, x - stale, > - best, b - backup, p - purge,
                 w - unused-weight-only
 Origin codes  : i - IGP, e - EGP, ? - incomplete

===============================================================================
BGP IPv4 Routes
===============================================================================
Original Attributes

Network        : 172.16.5.1/32
Nexthop        : 192.0.2.5
Path Id        : None
From           : 192.0.2.7
Res. Protocol  : SR-POLICY              Res. Metric    : 0
Res. Nexthop   : 0.0.0.0 (SR-POLICY)
Local Pref.    : 100                    Interface Name : NotAvailable
Aggregator AS  : None                   Aggregator     : None
Atomic Aggr.   : Not Atomic             MED            : None
AIGP Metric    : None                   IGP Cost       : 0
Connector      : None
Community      : color:01:600
Cluster        : 192.0.2.7
Originator Id  : 192.0.2.5              Peer Router Id : 192.0.2.7
Fwd Class      : None                   Priority       : None
Origin         : IGP
Flags          : Used Valid Best In-RTM
Route Source   : Internal
AS-Path        : No As-Path
Route Tag      : 0
Neighbor-AS    : n/a
DB Orig Val    : NotFound               Final Orig Val : N/A
Source Class   : 0                      Dest Class     : 0
Add Paths Send : Default
RIB Priority   : Normal
Last Modified  : 00h02m14s

Modified Attributes
---snip---
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
Routes : 1
```

```
===============================================================================
```

Advertising SR policies into BGP

Before advertising SR policies into BGP, all previous static SR policy configuration is removed. The simulated controller acts as the source of BGP advertised SR policies, and when an SR OS router advertises SR policies into BGP they must first be statically configured to provide the relevant information to populate the BGP path attributes. The following static SR policy is applied at the controller representing a similar SR policy to that previously configured at PE-1. The SR policy has a head-end of PE-1 (192.0.2.1), an endpoint of PE-5 (192.0.2.5), and a color of 600. The segment list is modified slightly to represent a list of strict hops using Adj-SIDs along the path PE-1-PE-2-PE-4-PE-6-PE-5.

```
# on controller:
configure {
    router "Base" {
        segment-routing {
            sr-policies {
                static-policy "color600-PE-1-PE-5" {
                    admin-state enable
                    color 600
                    endpoint 192.0.2.5
                    head-end 192.0.2.1
                    binding-sid 100002
                    distinguisher 600001005
                    segment-list 1 {
                        admin-state enable
                        segment 1 {
                            mpls-label 150012   # adj-SID int-PE-1-PE-2
                        }
                        segment 2 {
                            mpls-label 150024   # adj-SID int-PE-2-PE-4
                        }
                        segment 3 {
                            mpls-label 150046   # adj-SID int-PE-4-PE-6
                        }
                        segment 4 {
                            mpls-label 150065   # adj-SID int-PE-6-PE-5
                        }
                    }
                }
            }
        }
    }
}
```

To advertise the preceding SR policy into BGP, two steps are required at the controller. First, the **sr-policy-import** command must be configured under the **bgp** context. This command instructs BGP to import all statically configured non-local SR policies from the SR database into the BGP RIB, such that they can be advertised toward BGP peers supporting the SR policy address family. Second, a BGP peering is established with the Route Reflector RR-7 (192.0.2.7) for the address family, using the keyword **sr-policy-ipv4**. Although not shown, the relevant configuration is made on all routers for RR-7 to peer to all its clients for the same address family.

```
# on controller:
configure {
    router "Base" {
        bgp {
```

```
            sr-policy-import true # import non-local static SR policies into BGP RIB
            group "sr-policy" {
                peer-as 64496
                family {
                    sr-policy-ipv4 true
                }
            }
            neighbor "192.0.2.7" {
                group "sr-policy"
            }
        }
    }
}
```

Note:
SR OS Release 25.10.R2 supports propagation of SR policy routes across internal BGP peers. SR policy routes are not advertised to external BGP peers.

The following output shows the BGP RIB-Out for the SR policy address family at the controller and shows the SR policy advertised to RR-7 (192.0.2.7). The presence of an IPv4 address-specific Route-Target Extended Community encoding the head-end PE-1 address (192.0.2.1) allows for potential constraining of route propagation if required.

```
[/]
A:admin@PCE# show router bgp routes sr-policy-ipv4 hunt
===============================================================================
 BGP Router ID:192.0.2.254      AS:64496       Local AS:64496
===============================================================================
 Legend -
 Status codes  : u - used, s - suppressed, h - history, d - decayed, * - valid
                 l - leaked, x - stale, > - best, b - backup, p - purge,
                 w - unused-weight-only
 Origin codes  : i - IGP, e - EGP, ? - incomplete

===============================================================================
BGP SR-POLICY-v4 Routes
===============================================================================
-------------------------------------------------------------------------------
RIB In Entries
-------------------------------------------------------------------------------
RD/Color/End Pt: 600001005/600/192.0.2.5
BSID/Pref/TunnType: 100002/100/sr-policy
Nexthop        : 0.0.0.0
From           : BGP
Res. Nexthop   : n/a
Local Pref.    : None                  Interface Name : NotAvailable
Aggregator AS  : None                  Aggregator     : None
Atomic Aggr.   : Not Atomic            MED            : None
AIGP Metric    : None                  IGP Cost       : 0
Connector      : None
Community      : target:192.0.2.1:0
Cluster        : No Cluster Members
Originator Id  : None                  Peer Router Id : 0.0.0.0
Origin         : IGP
Flags          : Used Valid Best
Route Source   : Internal
AS-Path        : No As-Path
Route Tag      : 0
Neighbor-AS    : n/a
DB Orig Val    : N/A                   Final Orig Val : N/A
Source Class   : 0                     Dest Class     : 0
Add Paths Send : Default
```

```
Last Modified  : 00h00m31s

-------------------------------------------------------------------------------
RIB Out Entries
-------------------------------------------------------------------------------
RD/Color/End Pt: 600001005/600/192.0.2.5
BSID/Pref/TunnType: 100002/100/sr-policy
Nexthop        : 192.0.2.254
To             : 192.0.2.7
Res. Nexthop   : n/a
Local Pref.    : 100                    Interface Name : NotAvailable
Aggregator AS  : None                   Aggregator     : None
Atomic Aggr.   : Not Atomic             MED            : None
AIGP Metric    : None                   IGP Cost       : n/a
Connector      : None
Community      : target:192.0.2.1:0
Cluster        : No Cluster Members
Originator Id  : None                   Peer Router Id : 192.0.2.7
Origin         : IGP
AS-Path        : No As-Path
Route Tag      : 0
Neighbor-AS    : n/a
DB Orig Val    : N/A                    Final Orig Val : N/A
Source Class   : 0                      Dest Class     : 0

-------------------------------------------------------------------------------
Routes : 2
===============================================================================
```

The following output from PE-1 shows that the SR policy is active and that the first SID in the segment list has been correctly resolved; the owner is bgp.

```
[/]
A:admin@PE-1# show router segment-routing sr-policies bgp color 600 end-point 192.0.2.5

===============================================================================
SR-Policies Path
===============================================================================
-------------------------------------------------------------------------------
Type            : srMpls
Active          : Yes                   Owner          : bgp
Operational     : Yes
Color           : 600
Head            : 0.0.0.0               Endpoint Addr  : 192.0.2.5
RD              : 600001005             Preference     : 100
BSID            : 100002
TunnelId        : 917508                Age            : 60
Origin ASN      : 64496                 Origin         : 192.0.2.254
NumReEval       : 0                     LastReEvalReason: none
NumActPathChange: 0                     Last Change    : 01/29/2026 22:08:23
Maintenance Plcy:
Ret Path BFD Lbl:

Path Segment Lists:
Segment-List    : 1                     Weight         : 1
Num Segments    : 4                     Last Change    : 01/29/2026 21:45:49
  1 MPLS Label  : 150012                State          : resolved-up
  2 MPLS Label  : 150024                State          : N/A
  3 MPLS Label  : 150046                State          : N/A
  4 MPLS Label  : 150065                State          : N/A

===============================================================================
```

Verification is also made at PE-1 that the SR policy is correctly populated in the tunnel table.

```
[/]
A:admin@PE-1# show router tunnel-table 192.0.2.5 protocol sr-policy

===============================================================================
IPv4 Tunnel Table (Router: Base)
===============================================================================
Destination           Owner     Encap TunnelId  Pref   Nexthop        Metric
   Color
-------------------------------------------------------------------------------
192.0.2.5/32          sr-policy MPLS  917508    14     192.168.12.2   0
   600
-------------------------------------------------------------------------------
---snip---
===============================================================================
```

The procedure for traffic steering using an SR policy learned through BGP SR TE policy is the same as traffic steering using a statically configured SR policy, and is therefore not repeated here.

BSID anchor

The statically configured and BGP advertised SR policies used so far in this chapter have been instantiated on a head-end that uses the Color Extended Community to steer traffic. An alternative method of steering traffic using an SR policy is through the use of the BSID. If an incoming packet has an active label that matches the BSID of a valid SR policy, the packet is forwarded using that SR policy and the incoming label is swapped for the labels that the SR policy contains.

Using a BSID in this way is considered useful at domain interconnects such as ABRs or ASBRs. It provides opacity between the domains and protects the churn from one domain from entering another domain. In large networks, it has the additional benefit of reducing the number of labels an ingress router needs to impose, because the BSID can expand a single incoming SID/label stack (the BSID) into a much larger outgoing SID/label stack.

The example topology in Figure 7: Example topology is entirely IS-IS Level 2, so not constructed of multiple domains. However, it is still sufficient to show the use of BSID traffic steering. In the following example, PE-3 becomes a BSID anchor for an SR policy path extended between PE-1 and PE-5. This requires the instantiation of two SR policies:

- An SR policy at PE-3 with a segment list that constructs the required path to PE-5. Like every SR policy, it requires a BSID, but in this case the BSID is programmed in the Incoming Label Map (ILM) table and has a next-hop Label Forwarding Entry (NHLFE) that includes the segments (labels) in the segment list.
- An SR policy at PE-1 with a segment list specifying a path to PE-3, followed by a segment that references the relevant BSID programmed at PE-3.

The following output shows the SR policy advertised in BGP to PE-3. It uses color 700 and has an endpoint of PE-5 (192.0.2.5). Since traffic steering at PE-3 using the SR policy is achieved using the BSID, any color value could be used (although different colors may be needed to represent different path characteristics). Packets are classified upstream of PE-3 at PE-1, and the result of that classification selects the relevant BSID to meet the path requirements. The segment list programs a path to PE-5 using Adj-SIDs along the path PE-3-PE-4-PE-6-PE-5. The BSID value is 100001.

```
[/]
A:admin@PE-3# show router segment-routing sr-policies bgp color 700

===============================================================================
```

```
SR-Policies Path
===============================================================================
-------------------------------------------------------------------------------
Type            : srMpls
Active          : Yes                   Owner           : bgp
Operational     : Yes
Color           : 700
Head            : 0.0.0.0               Endpoint Addr   : 192.0.2.5
RD              : 700003005             Preference      : 100
BSID            : 100001
TunnelId        : 917506                Age             : 12
Origin ASN      : 64496                 Origin          : 192.0.2.254
NumReEval       : 0                     LastReEvalReason: none
NumActPathChange: 0                     Last Change     : 01/29/2026 22:10:27
Maintenance Plcy:
Ret Path BFD Lbl:

Path Segment Lists:
Segment-List    : 1                     Weight          : 1
Num Segments    : 3                     Last Change     : 01/29/2026 21:45:59
  1 MPLS Label  : 150034                State           : resolved-up
  2 MPLS Label  : 150046                State           : N/A
  3 MPLS Label  : 150065                State           : N/A

===============================================================================
```

The following output shows the SR policy advertised in BGP to PE-1. It uses color 700 and has an endpoint of PE-5 (192.0.2.5). The segment list programs a path that contains the following:

- The Node-SID of PE-3 (50403)
- The BSID programmed at PE-3 for the path to PE-5 (100001). When PE-3 pops its Node-SID and this label is exposed at PE-3, it swaps label 100001 for the label stack contained in the SR policy of that BSID.
- The Node-SID of PE-5 (50405)

```
[/]
A:admin@PE-1# show router segment-routing sr-policies bgp color 700

===============================================================================
SR-Policies Path
===============================================================================
-------------------------------------------------------------------------------
Type            : srMpls
Active          : Yes                   Owner           : bgp
Operational     : Yes
Color           : 700
Head            : 0.0.0.0               Endpoint Addr   : 192.0.2.5
RD              : 700001003             Preference      : 100
BSID            : 100003
TunnelId        : 917509                Age             : 9
Origin ASN      : 64496                 Origin          : 192.0.2.254
NumReEval       : 0                     LastReEvalReason: none
NumActPathChange: 0                     Last Change     : 01/29/2026 22:10:27
Maintenance Plcy:
Ret Path BFD Lbl:

Path Segment Lists:
Segment-List    : 1                     Weight          : 1
Num Segments    : 3                     Last Change     : 01/29/2026 21:45:49
  1 MPLS Label  : 50403                 State           : resolved-up
  2 MPLS Label  : 100001                State           : N/A
```

```
  3 MPLS Label  : 50405                State          : N/A

===============================================================================
```

A VPRN service (700) is extended between PE-1 and PE-5 with import/export Route-Target 64496:700, and with the auto-bind-tunnel resolution-filter set to SR policy at PE-1 (the complete VPRN service configuration is not shown for conciseness).

```
# on PE-1:
configure {
    service {
        vprn "VPRN_700" {
            admin-state enable
            service-id 700
            customer "1"
            autonomous-system 64496
            bgp-ipvpn {
                mpls {
                    admin-state enable
                    route-distinguisher "64496:700"
                    vrf-import {
                        policy ["vrf700-import"]
                    }
                    vrf-export {
                        policy ["vrf700-export"]
                    }
                    auto-bind-tunnel {
                        resolution filter
                        resolution-filter {
                            sr-policy true
                        }
                    }
                }
            }
        }
    }
}
```

PE-5 advertises prefix 172.31.5.1/32 as a VPN-IPv4 route. In addition to the Route-Target Extended Community attached to the VPN-IPv4 route, PE-5 also attaches a Color Extended Community with value 700. PE-5's VRF export policy is as follows:

```
# on PE-5:
configure {
    policy-options {
        community "vrf700-export" {
            member "target:64496:700" { }
        }
        community "vrf700-sr-policy" {
            member "color:00:700" { }
        }
        prefix-list "vrf700-prefixes" {
            prefix 172.31.5.1/32 type exact {
            }
        }
        policy-statement "vrf700-export" {
            entry 10 {
                from {
                    prefix-list ["vrf700-prefixes"]
                }
                to {
                    protocol {
```

```
                        name [bgp-vpn]
                    }
                }
                action {
                    action-type accept
                    community {
                        add ["vrf700-export" "vrf700-sr-policy"]
                    }
                }
            }
        }
    }
}
```

PE-1 also advertises prefix 172.31.1.1/32 as a VPN-IPv4 route to allow for connectivity to be validated end-to-end through both SR policies. The first of the following outputs shows PE-1s tunnel-table containing the advertised SR policy to PE-5 with tunnel ID 917509 and next-hop 192.0.2.3.

```
[/]
A:admin@PE-1# show router tunnel-table 192.0.2.5 protocol sr-policy

===============================================================================
IPv4 Tunnel Table (Router: Base)
===============================================================================
Destination           Owner     Encap TunnelId  Pref   Nexthop        Metric
   Color
-------------------------------------------------------------------------------
192.0.2.5/32          sr-policy MPLS  917508    14     192.168.12.2   0
   600
192.0.2.5/32          sr-policy MPLS  917509    14     192.0.2.3      0
   700
-------------------------------------------------------------------------------
---snip---
===============================================================================
```

The next output shows the route-table of VPRN 700 at PE-1 where the VPN-IPv4 prefix 172.31.5.1/32 advertised by PE-5 is resolved to an SR policy with tunnel ID 917509. As in the previous output, this is the SR policy advertised in BGP containing the BSID at PE-3.

```
[/]
A:admin@PE-1# show router 700 route-table

===============================================================================
Route Table (Service: 700)
===============================================================================
Dest Prefix[Flags]                              Type    Proto     Age        Pref
      Next Hop[Interface Name]                                      Metric
-------------------------------------------------------------------------------
172.31.1.1/32                                   Local   Local     00h00m46s  0
       loopback0-700                                                 0
172.31.5.1/32                                   Remote  BGP VPN   00h00m31s  170
       192.0.2.5 (tunneled:SR-Policy:917509)                         0
-------------------------------------------------------------------------------
No. of Routes: 2
---snip---
===============================================================================
```

The datapath between PE-1 and PE-5 is verified using a ping:

```
[/]
A:admin@PE-1# ping 172.31.5.1 router-instance "VPRN_700" source-address 172.31.1.1
```

```
PING 172.31.5.1 56 data bytes
64 bytes from 172.31.5.1: icmp_seq=1 ttl=64 time=3.36ms.
64 bytes from 172.31.5.1: icmp_seq=2 ttl=64 time=3.01ms.
. 64 bytes from 172.31.5.1: icmp_seq=3 ttl=64 time=3.01ms.
. 64 bytes from 172.31.5.1: icmp_seq=4 ttl=64 time=3.00ms.
. 64 bytes from 172.31.5.1: icmp_seq=5 ttl=64 time=3.43ms.

---- 172.31.5.1 PING Statistics ----
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min = 3.00ms, avg = 3.16ms, max = 3.43ms, stddev = 0.190ms
```

By enabling egress statistics for SR policies at PE-3, it is also possible to see the number of packets and octets being forwarded using the SR policy.

```
# on PE-3:
configure {
    router "Base" {
        segment-routing {
            sr-policies {
                egress-statistics {
                    admin-state enable
                }
            }
        }
    }
}

[/]
A:admin@PE-3# show router segment-routing sr-policies egress-statistics color 700 end-point
 192.0.2.5

===============================================================================
SR-Policies Egress Statistics
===============================================================================

Egress Statistics:

Color           : 700                   Endpoint Addr   : 192.0.2.5
Segment-List    : 1
TunnelId        : 917506                BSID            : 100001
Pkt Count       : 11                    Octet Count     : 1342

===============================================================================
```

Weighted Equal Cost Multipath

Support for weighted Equal Cost Multipath (ECMP) is provided with SR policies using multiple segment lists. Each segment list contains a path from the head-end to the endpoint, and each segment list contains a weight used to influence ECMP forwarding. The following output at the controller shows the use of multiple segment lists for an SR policy with a head-end of PE-1 (192.0.2.1), an endpoint of PE-6 (192.0.2.6), and a color of 800. Segment list 1 encodes a path consisting of Node-SIDs along the path PE-1-PE-3-PE-5-PE-6 and has a weight of 40. Segment list 2 encodes a path consisting of Node-SIDs along the path PE-1-PE-2-PE-4-PE-6 and has a weight of 60.

```
# on controller:
configure {
    router "Base" {
        segment-routing {
```

```
            sr-policies {
                static-policy "color800-PE-1-PE-6" {
                    admin-state enable
                    color 800
                    endpoint 192.0.2.6
                    head-end 192.0.2.1
                    binding-sid 100001
                    distinguisher 800001006
                    segment-list 1 {
                        admin-state enable
                        weight 40
                        segment 1 {
                            mpls-label 50403    # node-SID PE-3
                        }
                        segment 2 {
                            mpls-label 50405    # node-SID PE-5
                        }
                        segment 3 {
                            mpls-label 50406    # node-SID PE-6
                        }
                    }
                    segment-list 2 {
                        admin-state enable
                        weight 60
                        segment 1 {
                            mpls-label 50402    # node-SID PE-2
                        }
                        segment 2 {
                            mpls-label 50404    # node-SID PE-4
                        }
                        segment 3 {
                            mpls-label 50406    # node-SID PE-6
                        }
                    }
                }
            }
        }
    }
}
```

The SR policy at PE-1 shows that the SR policy is active and that both segment lists are resolved and up.

```
[/]
A:admin@PE-1# show router segment-routing sr-policies bgp color 800

===============================================================================
SR-Policies Path
===============================================================================
-------------------------------------------------------------------------------
Type            : srMpls
Active          : Yes                   Owner           : bgp
Operational     : Yes
Color           : 800
Head            : 0.0.0.0               Endpoint Addr   : 192.0.2.6
RD              : 800001006             Preference      : 100
BSID            : 100001
TunnelId        : 917510                Age             : 10
Origin ASN      : 64496                 Origin          : 192.0.2.254
NumReEval       : 0                     LastReEvalReason: none
NumActPathChange: 0                     Last Change     : 01/29/2026 22:14:04
Maintenance Plcy:
Ret Path BFD Lbl:
```

```
Path Segment Lists:
Segment-List    : 1                     Weight          : 40
Num Segments    : 3                     Last Change     : 01/29/2026 21:45:49
  1 MPLS Label  : 50403                 State           : resolved-up
  2 MPLS Label  : 50405                 State           : N/A
  3 MPLS Label  : 50406                 State           : N/A

Segment-List    : 2                     Weight          : 60
Num Segments    : 3                     Last Change     : 01/29/2026 21:45:49
  1 MPLS Label  : 50402                 State           : resolved-up
  2 MPLS Label  : 50404                 State           : N/A
  3 MPLS Label  : 50406                 State           : N/A

===============================================================================
```

The tunnel table at PE-1 shows two tunnels to PE-6 (192.0.2.6) owned by SR policy. Both entries reference the same tunnel ID of 917510, but the first entry has a next-hop of PE-3 (192.0.2.3) while the second entry has a next-hop of PE-2 (192.0.2.2).

```
[/]
A:admin@PE-1# show router tunnel-table 192.0.2.6 protocol sr-policy

===============================================================================
IPv4 Tunnel Table (Router: Base)
===============================================================================
Destination           Owner     Encap TunnelId  Pref   Nexthop        Metric
   Color
-------------------------------------------------------------------------------
192.0.2.6/32          sr-policy MPLS  917510    14     192.0.2.3      0
   800
192.0.2.6/32          sr-policy MPLS  917510    14     192.0.2.2      0
   800
-------------------------------------------------------------------------------
---snip---
===============================================================================
```

Conclusion

SR policies provide an effective way for instantiating traffic engineered SR tunnels that may be statically configured or advertised into BGP from either a controller or a router. Segments of paths constructed using SR policies can be loose or strict, using any combination of SIDs. The use of BSIDs also provides a way to interconnect domains and reduces the label stack imposition required at ingress routers. BGP can be used to advertise and instantiate SR policies that can be used as a method of steering traffic.

Flexible SR-TE Label Stack Allocation for BGP Services

This chapter provides information about the flexible segment routing traffic engineered (SR-TE) label stack allocation for BGP services.

Topics in this chapter include:

Applicability

The information and MD-CLI configuration in this chapter are based on SR OS Release 23.10.R3.

Flexible SR-TE label stack allocation for BGP services is supported on FP4-based platforms in SR OS Release 22.2.R1, and later.

Overview

FP4-based platforms support pushing 12 labels for SR-TE LSPs used in VPRN services and 10 labels for SR-TE LSPs used in EVPN VPLS or EVPN VPWS services. Layer 2 services, such as VPLS or Epipe services require an inner Ethernet header which reduces the space available for the label stack. In the case of R-VPLS or B-VPLS, even less label space is available for the label stack. For R-VPLS, only 9 labels can be pushed, and for B-VPLS, only 6.

Each SR-TE LSP is configured with a maximum transport label stack size, which is the sum of the values for the label stack size and the number of additional FRR labels. The label stack size ranges from 1 to 11 and the number of additional FRR labels ranges from 0 to 4, as follows. By default, the value for the label stack size is 6 and the number of additional FRR labels is 1.

```
[ex:/configure router "Base" mpls lsp "to-PE-2-empty" max-sr-labels]
A:admin@PE-1# label-stack-size ?

 label-stack-size <number>
 <number>        - <1..11>
 Dynamic Default - 6

    Maximum label stack size


[ex:/configure router "Base" mpls lsp "to-PE-2-empty" max-sr-labels]
A:admin@PE-1# additional-frr-labels ?

 additional-frr-labels <number>
```

```
 <number> - <0..4>
 Default  - 1

    Value for the maximum additional overhead labels
```

For VPRN services, the ingress label edge router (ILER) can push 12 labels, including the service label. A maximum of 11 labels remain for the sum of the label stack size and the number of additional FRR labels. The following error is raised because the sum of 10 and 2 exceeds the maximum of 11 labels:

```
*[ex:/configure router "Base" mpls lsp "to-PE-2-empty"]
A:admin@PE-1# commit
MINOR: MGMT_CORE #3001: configure router "Base" mpls lsp "to-PE-2-empty" max-sr-labels
label-stack-size
- Sum of label-stack-size and additional-frr-labels exceeds max labels per stack 11
```

Default egress label stack allocation for BGP services

In addition to the transport labels, the ILER pushes other fields into the packet, such as the service label, control word, Ethernet segment identifier (ESI), OAM labels, hash label, or entropy labels. The service label is present for all types of services, while the other labels depend on the service type. Some of these labels are always computed in the label stack allocation, such as the service label, OAM label, ESI, and control word. The label space for these labels is always reserved when computing the maximum stack; even when no Ethernet segment is used, one label is reserved for an ESI (for the purpose of BGP next hop resolution). Other labels are optional, such as the hash label or the entropy label, which are mutually exclusive. For the entropy label, two labels are required: the entropy label indicator (ELI) and the entropy label (EL) itself. These optional labels are only allocated when the hash or entropy label is configured. Table 2: Default egress label stack limits for BGP services shows an overview of the default egress label stack limits for BGP services, such as IP-VPN, EVPN-IFL, EVPN VPWS or EVPN VPLS, EVPN-IFF, and EVPN B-VPLS.

Table 2: Default egress label stack limits for BGP services

Features that reduce the label stack	IP-VPN (VPRN)	EVPN-IFL (VPRN)	EVPN VPLS or EVPN VPWS	EVPN-IFF (R-VPLS)	EVPN B-VPLS (PBB-EVPN)
Service label (always allocated)	1	1	1	1	1
OAM label (always allocated)	1	1	0	0	0
Control word (always allocated)	0	0	1	1	1
ESI label (always allocated)	0	0	1	0	0
Hash label (only when configured; mutually exclusive with EL)	1	1	0	0	0
Entropy EL+ELI (only when configured; mutually exclusive with the hash label)	2	2	2	2	2
Number of always-allocated labels	2	2	3	2	2

Features that reduce the label stack	IP-VPN (VPRN)	EVPN-IFL (VPRN)	EVPN VPLS or EVPN VPWS	EVPN-IFF (R-VPLS)	EVPN B-VPLS (PBB-EVPN)
Number of allocated labels, including the options	4	4	5	4	4
Number of available labels	12	12	10	9	6
Number of available transport labels when no options are configured	**10**	**10**	**7**	**7**	**4**
Number of available transport labels with options	**8**	**8**	**5**	**5**	**2**

The label space is always allocated for the first four features: the service label, OAM label, control word, and ESI label.

For IP-VPN or EVPN IFL VPRN services, only the service label and OAM label can be applied; therefore the transport label stack can have a maximum of 12 – 2 = 10 labels. In the case that the entropy label option is configured, two additional labels are reserved for EL and ELI and the transport label stack can have a maximum of 8 labels. When the hash label option is used, the number of available transport labels is only reduced by 1 and 9 transport labels are available.

For EVPN VPWS or EVPN VPLS services, the service label, control word, and ESI label are always allocated, even when the control word or ESI is not used. The number of available transport labels is 10 – 3 = 7 when no options are used. When the entropy label option is used, the number of available transport labels is further reduced by 2 and only 5 transport labels can be used.

For EVPN-IFF R-VPLS services, the traffic is either bridged traffic containing an ESI label or routed traffic without an ESI label. The ESI label is not accounted for because the routed encapsulation is always larger. For R-VPLS, the ILER can push a maximum of 9 labels, including the service label and control word. The number of available transport labels is 9 – 2 = 7, unless entropy labels are used. When the entropy label option is used, only 7 – 2 = 5 transport labels are available.

For B-VPLS services, the ILER can only push 6 labels, including the service label and control word. The number of available transport labels is only 6 – 2 = 4 when no options are used. When the entropy label option is used, only 4 – 2 = 2 transport labels are available.

SR-TE LSPs are configured with **max-sr-labels label-stack-size <..>** and **max-sr-labels additional-frr-labels <..>**. The number of these configured LSP labels must not exceed the number of available transport labels in the table. The BGP route next hop for the LSP cannot be resolved when the number of available labels is exceeded.

Dynamic egress label stack allocation for BGP services

With dynamic egress label stack allocation per service, SR OS accounts for the service label, but not for the labels that are not used. The **dynamic-egress-label-limit true** command extends the number of available labels in the dynamic egress label stack by not accounting for the labels that are not used. When the **dynamic-egress-label-limit** command is enabled, the OAM label cannot be used in VPRN services and **vprn-ping** and **vprn-trace** are not supported. In EVPN VPWS or EVPN VPLS services, the control word can still be used, as a result, the corresponding label space is allocated. In EVPN VPLS services, the ESI label is only accounted for if the VPLS has a SAP or SDP binding associated to an ES.

Dynamic egress label stack allocation is supported for IP-VPN, EVPN-IFL, EVPN VPWS, EVPN VPLS, EVPN-IFF, and EVPN B-VPLS. Table 3: Dynamic egress label stack limits for BGP services shows a comparison for the dynamic egress label stack allocation for IP-VPN, EVPN-IFL, EVPN VPLS, and EVPN VPWS, but a similar comparison can be made for EVPN-IFF and EVPN B-VPLS.

Table 3: Dynamic egress label stack limits for BGP services

Features that reduce the label stack	**dynamic-egress-label-limit false**			**dynamic-egress-label-limit true**		
	IP-VPN (VPRN)	**EVPN-IFL (VPRN)**	**EVPN VPLS EVPN Epipe**	**IP-VPN (VPRN)**	**EVPN-IFL (VPRN)**	**EVPN VPLS EVPN Epipe**
Service label	1 (always allocated)	1 (always allocated)	1 (always allocated)	1 (always allocated)	1 (always allocated)	1 (always allocated)
OAM label	1 (always allocated)	1 (always allocated)	0	0	0	0
Control word	0	0	1 (always allocated)	0	0	1
ESI label	0	0	1 (always allocated)	0	0	1
Hash label (mutually exclusive with EL)	1	1	0	1	0	0
Entropy EL+ELI (mutually exclusive with the hash label)	2	2	2	2	2	2
Number of always-allocated labels	2	2	3	1	1	1
Number of allocated labels, including the options	4	4	5	3	3	5
Number of available labels	12	12	10	12	12	10
Number of available transport labels when no options are configured	**10**	**10**	**7**	**11**	**11**	**9**
Number of available transport labels with options	**8**	**8**	**5**	**9**	**9**	**5**

When the **dynamic-egress-label-limit** command is enabled in an IP-VPN or EVPN-IFL service, the OAM label is not used and one additional transport label is available.

For IP-VPN or EVPN-IFL VPRN services, the ILER can push 12 labels and one of these labels is the service label. The remaining 11 labels are available for transport when no options are configured. When entropy labels are configured, 11 – 2 = 9 labels are available for transport.

For EVPN VPWS or EVPN VPLS services, the ILER can push 10 labels and one of these labels is the service label. The control word and ESI label are now considered as options. When no options are configured, 10 – 1 = 9 labels are available for the transport label stack. When the entropy label, control word, and ESI label are configured, only 5 labels are available for the transport label stack.

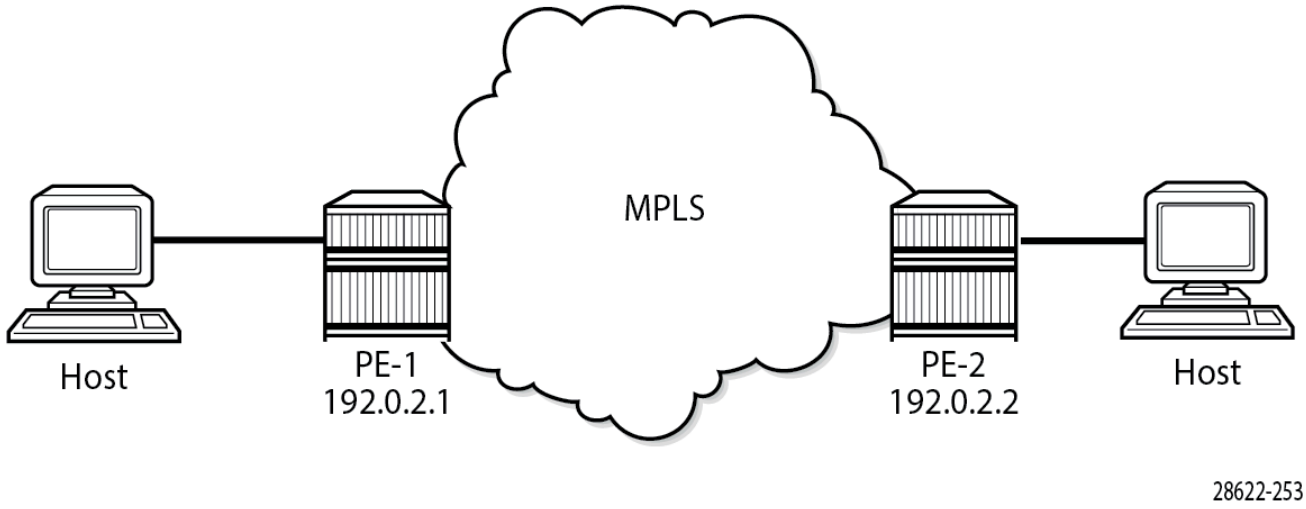
For EVPN-IFF services, the ILER can push 9 labels and one of these labels is the service label. When no options are configured, 9 – 1 = 8 labels are available for the transport label stack. When the control word and entropy label are configured, only 8 – 3 = 5 transport labels are available.

For EVPN B-VPLS services, the ILER can push 6 labels, including the service label. When no options are configured, 6 – 1 = 5 labels are available for the transport label stack. When the control word and entropy label are configured, only 5 – 3 = 2 transport labels are available.

Configuration

The Figure 8: Example topology consists of two nodes in an MPLS network:

Figure 8: Example topology



The initial configuration includes:

- cards, MDAs, and ports
- router interfaces
- SR-ISIS
- MPLS and RSVP enabled

BGP is configured for the EVPN and VPN-IPv4 address families; on PE-1 as follows:

```
#  on PE-1:
configure {
    router "Base" {
        autonomous-system 64500
        bgp {
            rapid-withdrawal true
            peer-ip-tracking true
```

```
            split-horizon true
            rapid-update {
                evpn true
            }
            group "internal" {
                peer-as 64500
                family {
                    vpn-ipv4 true
                    evpn true
                }
            }
            neighbor "192.0.2.2" {                          # on PE-2: 192.0.2.1
                group "internal"
            }
        }
```

The BGP configuration on PE-2 is similar.

SR-TE LSPs are configured between PE-1 and PE-2, as follows:

```
#  on PE-1:
configure {
    router "Base" {
        mpls {
            admin-state enable
            interface "int-PE-1-PE-2" {                     # on PE-2: "int-PE-2-PE-1"
            }
            path "empty" {
                admin-state enable
            }
            path "to-PE-2-strict" {                         # on PE-2: "to-PE-1-strict"
                admin-state enable
                hop 10 {
                    ip-address 192.168.12.2                 # on PE-2: 192.168.12.1
                    type strict
                }
            }
            lsp "to-PE-2-empty" {                           # on PE-2: "to-PE-1-empty"
                admin-state enable
                type p2p-sr-te
                to 192.0.2.2                                # on PE-2: to 192.0.2.1
                path-computation-method local-cspf
                max-sr-labels {                             # 10 labels - suited for VPRN
                    label-stack-size 8
                    additional-frr-labels 2
                }
                primary "empty" {
                }
            }
            lsp "to-PE-2-strict" {                          # on PE-2: "to-PE-1-strict"
                admin-state enable
                type p2p-sr-te
                to 192.0.2.2                                # on PE-2: to 192.0.2.1
                path-computation-method local-cspf
                max-sr-labels {                             # 10 labels - suited for VPRN
                    label-stack-size 8
                    additional-frr-labels 2
                }
                primary "to-PE-2-strict" {                  # on PE-2: "to-PE-1-strict"
                }
            }
        }
        rsvp {
            admin-state enable
```

```
            interface "int-PE-1-PE-2" {                    # on PE-2: "int-PE-2-PE-1"
            }
        }
```

The following services are described in this section:

- IP-VPN and EVPN-IFL services
- EVPN VPWS services
- EVPN VPLS services
- EVPN-IFF services
- EVPN PBB services

IP-VPN and EVPN-IFL services

For VPRN services, the ILER can push 12 labels: 1 service label, 1 OAM label, and 10 transport labels. VPRN-1 is an IP-VPN with the following configuration:

```
# on PE-1:
configure {
    service {
        vprn "VPRN-1" {
            admin-state enable
            description "VPRN-1 with BGP-IPVPN"
            service-id 1
            customer "1"
            bgp-ipvpn {
                mpls {
                    admin-state enable
                    route-distinguisher "192.0.2.1:1"      # on PE-2: 192.0.2.2:1
                    vrf-target {
                        community "target:64500:1"
                    }
                    auto-bind-tunnel {
                        resolution filter
                        resolution-filter {
                            bgp false
                            sr-te true
                        }
                    }
                }
            }
            interface "loopback" {
                loopback true
                ipv4 {
                    primary {
                        address 172.16.1.1                   # on PE-2: 172.16.1.2
                        prefix-length 32
                    }
                }
            }
        }
```

EVPN-IFL VPRN-2 is configured as follows:

```
# on PE-1:
configure {
    service {
```

```
        vprn "VPRN-2" {
            admin-state enable
            description "VPRN-2 with EVPN-IFL"
            service-id 2
            customer "1"
            bgp-evpn {
                mpls 1 {
                    admin-state enable
                    route-distinguisher "192.0.2.1:2"      # on PE-2: 192.0.2.2:2
                    vrf-target {
                        community "target:64500:2"
                    }
                    auto-bind-tunnel {
                        resolution filter
                        resolution-filter {
                            sr-te true
                        }
                    }
                }
            }
            interface "loopback" {
                loopback true
                ipv4 {
                    primary {
                        address 172.16.2.1                   # on PE-2: 172.16.2.2
                        prefix-length 32
                    }
                }
            }
        }
```

The BGP next hop is resolved for VPRN-1 and VPRN-2, as follows:

```
[/]
A:admin@PE-1# show router bgp next-hop vpn-ipv4 service-id 1
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================

===============================================================================
BGP VPN Next Hop
===============================================================================
VPN Next Hop                                                    Owner
   Autobind                                          FibProg    Reason
   Labels (User-labels)                              FlexAlgo   Metric
   Admin-tag-policy (strict-tunnel-tagging)                     Last Mod.
-------------------------------------------------------------------------------
192.0.2.2                                                       SR_TE
   sr-te                                             Y
   -- (2)                                            --         10
   -- (N)                                                       00h04m09s
-------------------------------------------------------------------------------
Next Hops : 1
===============================================================================

[/]
A:admin@PE-1# show router bgp next-hop evpn service-id 2
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================

===============================================================================
BGP VPN Next Hop
```

```
===============================================================================
VPN Next Hop                                                  Owner
   Autobind                                          FibProg  Reason
   Labels (User-labels)                              FlexAlgo Metric
   Admin-tag-policy (strict-tunnel-tagging)                   Last Mod.
-------------------------------------------------------------------------------
192.0.2.2                                                     SR_TE
   sr-te                                             Y
   -- (2)                                            --       10
   -- (N)                                                     00h04m09s
-------------------------------------------------------------------------------
Next Hops : 1
===============================================================================
```

The number of user labels is 2 and corresponds to the sum of the service label and OAM label.

The route table for VPRN-1 on PE-1 shows that the BGP VPN-IPv4 route to 172.16.1.2/32 uses an SR-TE tunnel with ID 655362 to PE-2, as follows:

```
[/]
A:admin@PE-1# show router service-name "VPRN-1" route-table

===============================================================================
Route Table (Service: 1)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
172.16.1.1/32                                 Local   Local     00h04m59s  0
       loopback                                                      0
172.16.1.2/32                                 Remote  BGP VPN   00h04m51s  170
       192.0.2.2 (tunneled:SR-TE:655362)                             10
-------------------------------------------------------------------------------
No. of Routes: 2
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
===============================================================================
```

The route table for VPRN-2 on PE-1 shows that the EVPN-IFL route to 172.16.2.2/32 also uses the SR-TE tunnel with ID 655362 to PE-2, as follows:

```
[/]
A:admin@PE-1# show router service-name "VPRN-2" route-table

===============================================================================
Route Table (Service: 2)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
172.16.2.1/32                                 Local   Local     00h04m59s  0
       loopback                                                      0
172.16.2.2/32                                 Remote  EVPN-IFL  00h04m51s  170
       192.0.2.2 (tunneled:SR-TE:655362)                             10
-------------------------------------------------------------------------------
No. of Routes: 2
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
```

```
       S = Sticky ECMP requested
===============================================================================
```

The following illustrates that the BGP next hop cannot be resolved in a VPRN when the number of labels configured in the SR-TE LSP exceeds 10. In this example, the label stack size is increased to 9, while the number of additional FRR labels remains 2, as follows:

```
#  on PE-1:
configure {
    router "Base" {
        mpls {
            lsp "to-PE-2-empty" {                  # on PE-2: "to-PE-1-empty"
                max-sr-labels {
                    label-stack-size 9
                    additional-frr-labels 2
                }
            }
            lsp "to-PE-2-strict"                   # on PE-2: "to-PE-1-strict"
                max-sr-labels {
                    label-stack-size 9
                    additional-frr-labels 2
                }
            }
```

The following shows that the BGP next hop cannot be resolved for the IP-VPN VPRN-1:

```
[/]
A:admin@PE-1# show router bgp next-hop vpn-ipv4
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================

===============================================================================
BGP VPN Next Hop
===============================================================================
VPN Next Hop                                                Owner
   Autobind                                        FibProg  Reason
   Labels (User-labels)                            FlexAlgo Metric
   Admin-tag-policy (strict-tunnel-tagging)                 Last Mod.
-------------------------------------------------------------------------------
192.0.2.2                                                   --
   sr-te                                           N        LabelStackLimit
   -- (2)                                          --
   -- (N)                                                   00h00m14s
-------------------------------------------------------------------------------
Next Hops : 1
===============================================================================
```

The route is not programmed in the FIB because of a LabelStackLimit error.

In a similar way, the BGP next hop cannot be resolved for the EPVN-IFL VPRN-2, as follows:

```
[/]
A:admin@PE-1# show router bgp next-hop evpn service-id 2
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================

===============================================================================
BGP VPN Next Hop
===============================================================================
VPN Next Hop                                                Owner
```

```
   Autobind                                        FibProg  Reason
   Labels (User-labels)                            FlexAlgo Metric
   Admin-tag-policy (strict-tunnel-tagging)                 Last Mod.
-------------------------------------------------------------------------------
192.0.2.2                                                   --
   sr-te                                           N        LabelStackLimit
   -- (2)                                          --
   -- (N)                                                   00h00m14s
-------------------------------------------------------------------------------
Next Hops : 1
===============================================================================
```

The route table for VPRN-1 does not contain a route to 172.16.1.2/32 anymore; only the local route remains, as follows:

```
[/]
A:admin@PE-1# show router service-name "VPRN-1" route-table

===============================================================================
Route Table (Service: 1)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
172.16.1.1/32                                 Local   Local     00h05m52s  0
       loopback                                                      0
-------------------------------------------------------------------------------
No. of Routes: 1
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
===============================================================================
```

In a similar way, the route table for VPRN-2 contains only the local route, as follows:

```
[/]
A:admin@PE-1# show router service-name "VPRN-2" route-table

===============================================================================
Route Table (Service: 2)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
172.16.2.1/32                                 Local   Local     00h05m52s  0
       loopback                                                      0
-------------------------------------------------------------------------------
No. of Routes: 1
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
===============================================================================
```

If the OAM label is not required, the **dynamic-egress-label-limit** command can be enabled in the VPRN services, as follows:

```
# on PE-1, PE-2:
configure {
    service {
```

```
        vprn "VPRN-1" {
            bgp-ipvpn {
                mpls {
                    dynamic-egress-label-limit true
                }
            }
        }
        vprn "VPRN-2" {
            bgp-evpn {
                mpls 1 {
                    dynamic-egress-label-limit true
                }
            }
        }
```

The ILER can push 12 labels: 1 service label and 11 transport labels. The BGP next hop can be resolved for VPRN-1, as follows:

```
[/]
A:admin@PE-1# show router bgp next-hop vpn-ipv4 service-id 1
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================

===============================================================================
BGP VPN Next Hop
===============================================================================
VPN Next Hop                                                  Owner
   Autobind                                          FibProg  Reason
   Labels (User-labels)                              FlexAlgo Metric
   Admin-tag-policy (strict-tunnel-tagging)                   Last Mod.
-------------------------------------------------------------------------------
192.0.2.2                                                     SR_TE
   sr-te                                             Y
   -- (1)                                            --       10
   -- (N)                                                     00h00m10s
-------------------------------------------------------------------------------
Next Hops : 1
===============================================================================
```

The number of user labels is 1 and only the service label is accounted for.

In a similar way, the BGP next hop can be resolved for VPRN-2 and the number of user labels is 1, as follows:

```
[/]
A:admin@PE-1# show router bgp next-hop evpn service-id 2
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================

===============================================================================
BGP VPN Next Hop
===============================================================================
VPN Next Hop                                                  Owner
   Autobind                                          FibProg  Reason
   Labels (User-labels)                              FlexAlgo Metric
   Admin-tag-policy (strict-tunnel-tagging)                   Last Mod.
-------------------------------------------------------------------------------
192.0.2.2                                                     SR_TE
   sr-te                                             Y
   -- (1)                                            --       10
```

```
   -- (N)                                                    00h00m10s
-------------------------------------------------------------------------------
Next Hops : 1
===============================================================================
```

The following shows that the **dynamic-egress-label-limit** command is enabled for IP-VPN VPRN-1:

```
[/]
A:admin@PE-1# show service id "VPRN-1" bgp-ipvpn

===============================================================================
Service 1 BGP-IPVPN MPLS Information
===============================================================================
Admin State       : Up                  Oper State        : Up
VRF Import        : None
VRF Export        : None
Route Dist.       : None
Oper Route Dist   : 192.0.2.1:1
Oper RD Type      : configured
Route Target      : target:64500:1
Route Target Impor: None
Route Target Expor: None
Domain-Id         : None
Dyn Egr Lbl Limit : Enabled

Auto-Bind Tunnel
Resolution        : filter              Strict Tnl Tag    : False
ECMP              : 1                   Flex Algo FB      : False
Weighted ECMP     : False
BGP Instance      : 1
Filter Tunnel Type: sr-te
===============================================================================
```

The following shows that the **dynamic-egress-label-limit** command is enabled for EVPN-IFL VPRN-2:

```
[/]
A:admin@PE-1# show service id "VPRN-2" bgp-evpn

===============================================================================
BGP EVPN MPLS Table
===============================================================================
Admin State        : Up                     Oper State     : Up
VRF Import         : None
VRF Export         : None
Route Dist.        : 192.0.2.1:2
Oper Route Dist.   : 192.0.2.1:2
Oper RD Type       : configured
Route Target       : target:64500:2
Route Target Import: None
Route Target Export: None
Default Route Tag  : None
Domain-Id          : None
Dyn Egr Lbl Limit  : Enabled
EVI                : 0

Advertise          : Disabled
Weighted ECMP      : Disabled

Auto-Bind Tunnel
Resolution         : filter                 Strict Tnl Tag : False
ECMP               : 1                      Flex Algo FB   : False
Bgp Instance       : 1
Filter Tunnel Types: sr-te
```

```
Tunnel Encap
MPLS              : True                  MPLSoUDP       : False
===============================================================================
```

EVPN VPWS services

Epipe-3 is configured as follows:

```
#  on PE-1:
configure {
    service {
        epipe "Epipe-3" {
            admin-state enable
            service-id 3
            customer "1"
            bgp 1 {
            }
            sap 1/1/c10/1:3 {
                description "SAP to CE-31"                # on PE-2: SAP to CE-32
            }
            bgp-evpn {
                evi 3
                local-attachment-circuit "ac-1" {         # on PE-2: ac-2
                    eth-tag 1                             # on PE-2: eth-tag 2
                }
                remote-attachment-circuit "ac-2" {        # on PE-2: ac-1
                    eth-tag 2                             # on PE-2: eth-tag 1
                }
                mpls 1 {
                    admin-state enable
                    auto-bind-tunnel {
                        resolution filter
                        resolution-filter {
                            sr-te true
                        }
                    }
                }
            }
        }
```

Layer 2 services, such as EVPN VPWS or EVPN VPLS services, cannot have the same number of labels in the stack as for VPRN services. A LabelStackLimit error causes Epipe-3 to be in an operationally down state, as follows:

```
[/]
A:admin@PE-1# show service service-using epipe

===============================================================================
Services [epipe]
===============================================================================
ServiceId    Type      Adm  Opr  CustomerId Service Name
-------------------------------------------------------------------------------
3            Epipe     Up   Down 1          Epipe-3
-------------------------------------------------------------------------------
Matching Services : 1
-------------------------------------------------------------------------------
===============================================================================
```

The following output shows that the BGP next hop cannot be resolved because of a LabelStackLimit error. By default, the number of user labels is 3 and accounts for 1 for the service label, 1 for the ESI label, and 1 for the control word, as follows:

```
[/]
A:admin@PE-1# show router bgp next-hop evpn service-id 3
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================

===============================================================================
BGP VPN Next Hop
===============================================================================
VPN Next Hop                                                 Owner
   Autobind                                        FibProg   Reason
   Labels (User-labels)                            FlexAlgo  Metric
   Admin-tag-policy (strict-tunnel-tagging)                  Last Mod.
-------------------------------------------------------------------------------
192.0.2.2                                                    --
   sr-te                                           N         LabelStackLimit
   -- (3)                                          --
   -- (N)                                                    00h04m24s
-------------------------------------------------------------------------------
Next Hops : 1
===============================================================================
```

For EVPN VPWS and EVPN VPLS services, the ILER can push 10 labels, including the service label, ESI label, and control word; therefore, the sum of the label stack size and the additional FRR labels cannot exceed 10 – 3 = 7 transport labels. However, when the service is configured with **dynamic-egress-label-limit true**, only the service label must be accounted for and the number of transport labels can be 10 – 1 = 9. The SR-TE LSPs are configured with a label stack size of 9 and without additional FRR labels; the Epipe service is configured with **dynamic-egress-label-limit true**, as follows:

```
# on PE-1:
configure {
    router "Base" {
        mpls {
            lsp "to-PE-2-empty" {                          # on PE-2: "to-PE-1-empty"
                max-sr-labels {
                    label-stack-size 9
                    additional-frr-labels 0
                }
            }
            lsp "to-PE-2-strict"                           # on PE-2: "to-PE-1-strict"
                max-sr-labels {
                    label-stack-size 9
                    additional-frr-labels 0
                }
            }
        }
    }
    service {
        epipe "Epipe-3" {
            bgp-evpn {
                mpls 1 {
                    dynamic-egress-label-limit true
                }
            }
```

The BGP next hop is now resolved and the route is programmed in the FIB. The number of user labels is 1 for the service label, as follows:

```
[/]
A:admin@PE-1# show router bgp next-hop evpn service-id 3
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================

===============================================================================
BGP VPN Next Hop
===============================================================================
VPN Next Hop                                                    Owner
   Autobind                                          FibProg    Reason
   Labels (User-labels)                              FlexAlgo   Metric
   Admin-tag-policy (strict-tunnel-tagging)                     Last Mod.
-------------------------------------------------------------------------------
192.0.2.2                                                       SR_TE
   sr-te                                             Y
   -- (1)                                            --         10
   -- (N)                                                       00h00m18s
-------------------------------------------------------------------------------
Next Hops : 1
===============================================================================
```

Epipe-3 is now in an operationally up state, as follows:

```
[/]
A:admin@PE-1# show service service-using epipe

===============================================================================
Services [epipe]
===============================================================================
ServiceId    Type      Adm  Opr  CustomerId Service Name
-------------------------------------------------------------------------------
3            Epipe     Up   Up   1          Epipe-3
-------------------------------------------------------------------------------
Matching Services : 1
-------------------------------------------------------------------------------
===============================================================================
```

From Epipe-3 on PE-1, the BGP EVPN-MPLS destination 192.0.2.2 can be reached via an SR-TE tunnel with ID 655362, as follows:

```
[/]
A:admin@PE-1# show service id "Epipe-3" evpn-mpls

===============================================================================
BGP EVPN-MPLS Dest (Instance 1)
===============================================================================
TEP Address                         Egr Label              Last Change
                                     Transport:Tnl-id
-------------------------------------------------------------------------------
192.0.2.2                           524286                 02/22/2024 14:13:49
                                    sr-te:655362
-------------------------------------------------------------------------------
Number of entries : 1
-------------------------------------------------------------------------------
===============================================================================
---snip---
```

EVPN VPLS services

On PE-1 and PE-2, EVPN VPLS "VPLS-4" is configured:

```
# on PE-1:
configure {
    service {
        vpls "VPLS-4" {
            admin-state enable
            service-id 4
            customer "1"
            bgp 1 {
            }
            bgp-evpn {
                evi 4
                mpls 1 {
                    admin-state enable
                    ingress-replication-bum-label true
                    ecmp 2
                    auto-bind-tunnel {
                        resolution filter
                        resolution-filter {
                            sr-te true
                        }
                    }
                }
            }
            sap 1/1/c10/1:4 {
                description "SAP to CE-41"        # on PE-2: "SAP to CE-42"
            }
        }
```

The maximum number of transport labels in EVPN VPLS services is the same as for EVPN VPWS services. By default, the ILER can push a maximum of 10 – 3 = 7 transport labels. The SR-TE LSPs are configured with a label stack size of 9 labels; therefore, a LabelStackLimit error occurs and the BGP next hop cannot be resolved. The number of user labels is 3 and accounts for 1 for the service label, 1 for the ESI label, and 1 for the control word, as follows:

```
[/]
A:admin@PE-1# show router bgp next-hop evpn service-id 4
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================

===============================================================================
BGP VPN Next Hop
===============================================================================
VPN Next Hop                                                  Owner
   Autobind                                        FibProg    Reason
   Labels (User-labels)                            FlexAlgo   Metric
   Admin-tag-policy (strict-tunnel-tagging)                   Last Mod.
-------------------------------------------------------------------------------
192.0.2.2                                                     --
   sr-te                                           N          LabelStackLimit
   -- (3)                                          --
   -- (N)                                                     00h00m11s
-------------------------------------------------------------------------------
Next Hops : 1
===============================================================================
```

The number of available transport labels can be increased when **dynamic-egress-label-limit** is enabled on PE-1 and PE-2, as follows:

```
# on PE-1, PE-2:
configure {
    service {
        vpls "VPLS-4" {
            bgp-evpn {
                mpls 1 {
                    dynamic-egress-label-limit true
```

With this configuration, only the service label is accounted for and the number of user labels is reduced to 1. The BGP next hop can be resolved, as follows:

```
[/]
A:admin@PE-1# show router bgp next-hop evpn service-id 4
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================

===============================================================================
BGP VPN Next Hop
===============================================================================
VPN Next Hop                                                  Owner
   Autobind                                          FibProg  Reason
   Labels (User-labels)                              FlexAlgo Metric
   Admin-tag-policy (strict-tunnel-tagging)                   Last Mod.
-------------------------------------------------------------------------------
192.0.2.2                                                     SR_TE
   sr-te                                             Y
   -- (1)                                            --       10
   -- (N)                                                     00h02m21s
-------------------------------------------------------------------------------
Next Hops : 1
===============================================================================
```

The following EVPN-MPLS destinations are established in VPLS-4 on PE-1:

```
[/]
A:admin@PE-1# show service id "VPLS-4" evpn-mpls

===============================================================================
BGP EVPN-MPLS Dest (Instance 1)
===============================================================================
TEP Address                     Transport:Tnl     Egr Label  Oper  Mcast  Num
                                                             State        MACs
-------------------------------------------------------------------------------
192.0.2.2                       sr-te:655362      524282     Up    bum    0
192.0.2.2                       sr-te:655362      524283     Up    none   1
-------------------------------------------------------------------------------
Number of entries: 2
-------------------------------------------------------------------------------
===============================================================================
---snip---
```

EVPN-IFF services

The R-VPLS configuration on PE-1 is as follows:

```
# on PE-1:
configure {
    service {
        vpls "R-VPLS-6" {
            admin-state enable
            description "R-VPLS 6 - broadcast domain 1"
            service-id 6
            customer "1"
            routed-vpls {
            }
            sap 1/1/c10/1:6 {
                description "SAP to CE-61"
            }
        }
        vprn "ip-vrf-66" {
            admin-state enable
            service-id 66
            customer "1"
            interface "int-R-VPLS-6" {           # toward BD 1 with local CEs
                mac 00:00:00:16:06:01
                ipv4 {
                    primary {
                        address 172.16.6.1
                        prefix-length 24
                    }
                    vrrp 1 {
                        backup [172.16.6.254]
                        passive true
                        ping-reply true
                        traceroute-reply true
                    }
                }
                vpls "R-VPLS-6" {
                }
            }
            interface "int-sbd-600" {            # toward PE-2
                mac 00:00:00:00:60:01
                ipv4 {
                    primary {
                        address 10.0.6.1
                        prefix-length 24
                    }
                }
                vpls "sbd-600" {
                }
            }
        }
        vpls "sbd-600" {
            admin-state enable
            description "supplementary broadcast domain R-VPLS 600 - backhaul"
            service-id 600
            customer "1"
            routed-vpls {
            }
            bgp 1 {
                route-distinguisher "192.0.2.1:600"
            }
            bgp-evpn {
```

```
                evi 600
                routes {
                    ip-prefix {
                        advertise true
                    }
                }
                mpls 1 {
                    admin-state enable
                    auto-bind-tunnel {
                        resolution filter
                        resolution-filter {
                            sr-te true
                        }
                    }
                }
            }
        }
```

The R-VPLS configuration on PE-2 is similar.

For R-VPLS services, the ILER can push a maximum of 9 labels, including the service label and control word. By default, a maximum of 7 transport labels can be pushed. When **dynamic-egress-label-limit** is enabled in the R-VPLS, a maximum of 8 transport labels are available. The SR-TE LSPs are configured with a label stack size of 8 labels and **dynamic-egress-label-limit** is enabled in R-VPLS "sbd-600", as follows:

```
# on PE-1:
configure {
    router "Base" {
        mpls {
            lsp "to-PE-2-empty" {                 # on PE-2: lsp "to-PE-1-empty"
                max-sr-labels {
                    label-stack-size 8
                    additional-frr-labels 0
                }
            }
            lsp "to-PE-2-strict" {                # on PE-2: lsp "to-PE-1-strict"
                max-sr-labels {
                    label-stack-size 8
                    additional-frr-labels 0
                }
            }
        }
    }
    service {
        vpls "sbd-600" {
            bgp-evpn {
                mpls 1 {
                    dynamic-egress-label-limit true
                }
            }
```

The BGP next hop can be resolved in R-VPLS "sbd-600", as follows:

```
[/]
A:admin@PE-1# show router bgp next-hop evpn service-id 600
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================

===============================================================================
```

```
BGP VPN Next Hop
===============================================================================
VPN Next Hop                                                  Owner
   Autobind                                          FibProg  Reason
   Labels (User-labels)                              FlexAlgo Metric
   Admin-tag-policy (strict-tunnel-tagging)                   Last Mod.
-------------------------------------------------------------------------------
192.0.2.2                                                     SR_TE
   sr-te                                             Y
   rvpls  (1)                                        --       10
   -- (N)                                                     00h00m18s
-------------------------------------------------------------------------------
Next Hops : 1
===============================================================================
```

The following EVPN-MPLS destination is established in R-VPLS "sbd-600":

```
[/]
A:admin@PE-1# show service id "sbd-600" evpn-mpls

===============================================================================
BGP EVPN-MPLS Dest (Instance 1)
===============================================================================
TEP Address                    Transport:Tnl     Egr Label  Oper  Mcast  Num
                                                            State        MACs
-------------------------------------------------------------------------------
192.0.2.2                      sr-te:655362      524281     Up    bum    1
-------------------------------------------------------------------------------
Number of entries: 1
-------------------------------------------------------------------------------
===============================================================================
---snip---
```

EVPN PBB services

The configuration for B-VPLS and I-VPLS is as follows:

```
# on PE-1:
configure {
     service {
         vpls "B-VPLS-500" {
            admin-state enable
            service-id 500
            customer "1"
            service-mtu 2000
            pbb-type b-vpls
            pbb {
                source-bmac {
                    address 00:00:00:00:00:01
                }
            }
            bgp 1 {
            }
            bgp-evpn {
                evi 500
                mpls 1 {
                    admin-state enable
                    auto-bind-tunnel {
                        resolution filter
                        resolution-filter {
```

```
                        sr-te true
                    }
                }
            }
        }
        vpls "I-VPLS-5" {
            admin-state enable
            service-id 5
            customer "1"
            pbb-type i-vpls
            pbb {
                backbone-vpls "B-VPLS-500" {
                    isid 5
                }
            }
            sap 1/1/c10/1:5 {
                description "SAP to CE-51"            # on PE-2: SAP to CE-52
            }
        }
```

The SR-TE LSPs are configured with a label stack size of 8 labels; therefore, a LabelStackLimit error occurs for the B-VPLS and the BGP next hop cannot be resolved, as follows:

```
[/]
A:admin@PE-1# show router bgp next-hop evpn service-id 500
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================

===============================================================================
BGP VPN Next Hop
===============================================================================
VPN Next Hop                                                    Owner
   Autobind                                          FibProg    Reason
   Labels (User-labels)                              FlexAlgo   Metric
   Admin-tag-policy (strict-tunnel-tagging)                     Last Mod.
-------------------------------------------------------------------------------
192.0.2.2                                                       --
   sr-te                                             N          LabelStackLimit
   bvpls  (2)                                        --
   -- (N)                                                       00h00m12s
-------------------------------------------------------------------------------
Next Hops : 1
===============================================================================
```

The LER can only push maximum 6 labels for the B-VPLS, including the service label and the control word. By default, only 4 transport labels are available when no options are used. When **dynamic-egress-label-limit** is enabled in the B-VPLS, only the service label is accounted for and maximum 5 labels are available for transport. The SR-TE LSPs are configured with a label stack size of 5 and **dynamic-egress-label-limit** is enabled in the B-VPLS:

```
# on PE-1:
configure {
    router "Base" {
        mpls {
            lsp "to-PE-2-empty" {                       # on PE-2: LSP "to-PE-1-empty"
                max-sr-labels {
                    label-stack-size 5
                    additional-frr-labels 0
                }
```

```
            }
            lsp "to-PE-2-strict" {                      # on PE-2: LSP "to-PE-1-strict"
                max-sr-labels {
                    label-stack-size 5
                    additional-frr-labels 0
                }
            }
        }
    }
    service {
        vpls "B-VPLS-500" {
            bgp-evpn {
                mpls 1 {
                    dynamic-egress-label-limit true
                }
            }
```

The BGP next hop can now be resolved for the B-VPLS, as follows:

```
[/]
A:admin@PE-1# show router bgp next-hop evpn service-id 500
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================

===============================================================================
BGP VPN Next Hop
===============================================================================
VPN Next Hop                                                Owner
   Autobind                                    FibProg      Reason
   Labels (User-labels)                        FlexAlgo     Metric
   Admin-tag-policy (strict-tunnel-tagging)                 Last Mod.
-------------------------------------------------------------------------------
192.0.2.2                                                   SR_TE
   sr-te                                       Y
   bvpls  (1)                                  --           10
   -- (N)                                                   00h00m13s
-------------------------------------------------------------------------------
Next Hops : 1
===============================================================================
```

In B-VPLS-500 on PE-1, the EVPN-MPLS destination 192.0.2.2 can be reached via an SR-TE tunnel with ID 655362, as follows:

```
[/]
A:admin@PE-1# show service id "B-VPLS-500" evpn-mpls

===============================================================================
BGP EVPN-MPLS Dest (Instance 1)
===============================================================================
TEP Address                     Transport:Tnl     Egr Label  Oper  Mcast  Num
                                                             State        MACs
-------------------------------------------------------------------------------
192.0.2.2                       sr-te:655362      524280     Up    bum    1
-------------------------------------------------------------------------------
Number of entries: 1
-------------------------------------------------------------------------------
===============================================================================
---snip---
```

Conclusion

Flexible SR-TE label stack allocation can increase the number of available transport labels in services when OAM labels, ESI labels, or control word are not used.

Inter-AS Model C VPRN Using MPLS Forwarding Policies and Segment Routing Policies

This chapter provides information about Inter-AS Model C VPRN using MPLS Forwarding Policies and Segment Routing Policies.

Topics in this chapter include:

Applicability

The information and configuration in this chapter are based on MD-CLI for SR OS Release 21.7.R1. MPLS label binding forwarding policies and segment routing policies are supported in SR OS Release 16.0.R1, or later. MPLS endpoint forwarding policies and ECMP are supported in SR OS Release 16.0.R4, or later.

Overview

In this configuration, MPLS forwarding policies are combined with segment routing policies. In the remainder of this chapter, SR refers to "Segment Routing", unless specified otherwise. Product and release references, such as 7750 SR and SR OS, continue to refer to "Service Router".

MPLS forwarding policies

SR OS uses the following table management to forward packets:

- Route Table Manager (RTM) for IP packets matching IP route prefixes in the Global Route Table (GRT) that resolve to IP next-hops or tunnel next-hops (for IGP shortcuts).
- Tunnel Table Manager (TTM) containing Next-Hop Label Forwarding Entries (NHLFEs) to forward IP packets for routes in GRT or VPRN using tunnels. The resolved next-hop of the IP packet is matched to the far-end address of the TTM entry.
- Incoming Label Map (ILM) containing labels matching a specific Forwarding Equivalence Class (FEC), such that packets with this label are sent to the destination of the FEC.

The ILM tunnel is programmed via the service module, the MPLS module, and various control plane protocols supporting labeled tunnels or FECs. The GRT and TTM provide some flexibility, but do not allow customization, such as the ability to create specific sets of IP direct next-hops, IP indirect next-hops, or tunnel next-hops for a specific set of flows or prefixes. For more flexibility, the following can be configured:

- static routes

- traffic steering; for example, using Openflow, and Policy-Based Routing (PBR)
- MPLS forwarding policies

MPLS forwarding policies establish Static Label Routes (SLRs). The binding label of the forwarding policy is popped when matched on an incoming packet. If no pushed label is configured, then it becomes a swap to implicit-null, which is essentially a pop operation. After the incoming label is popped, the exposed packet payload (or the next label after the top label is removed) is forwarded via the configured next-hop of the MPLS forwarding policy. The next-hop is looked up in the route table and can be direct or indirect. A direct next-hop is an attached local interface; an indirect next-hop is a resolved route.

MPLS label-binding forwarding policies use labels from a reserved label block also known as a Segment Routing Local Block (SRLB), whereas node SIDs in segment routing use the Segment Routing Global Block (SRGB) instead. An SRLB is used for the following:

- static adjacency SIDs
- adjacency set SIDs
- SR policy binding SIDs (BSIDs)
- MPLS forwarding policy binding labels

MPLS forwarding policies allow the forwarding of packets over a set of user-defined next-hops: either direct next-hops (with option to push a label stack) or indirect next-hops.

MPLS forwarding policies are validated as follows:

- the binding label must be in the label range of the defined reserved label block (SRLB) and it must be unused; the same label cannot be allocated more than once
- the direct next-hop interfaces must be up
- the indirect next-hops must be reachable

MPLS forwarding policies work in one of two modes:

- ILM mode: label binding policy for labeled packets
- LTN mode: endpoint policy for unlabeled packets (this is beyond the scope of this chapter)

The data model of a forwarding policy represents the primary and the backup next-hop as a Next-Hop Group (NHG) and models the ECMP as the set of NHGs. Flows of prefixes can be switched on a per-NHG basis -without disturbing flows forwarded over other NHGs of the policy- from the failing primary next-hop to the backup next-hop or from the backup next-hop to the restored primary next-hop.

SR policies

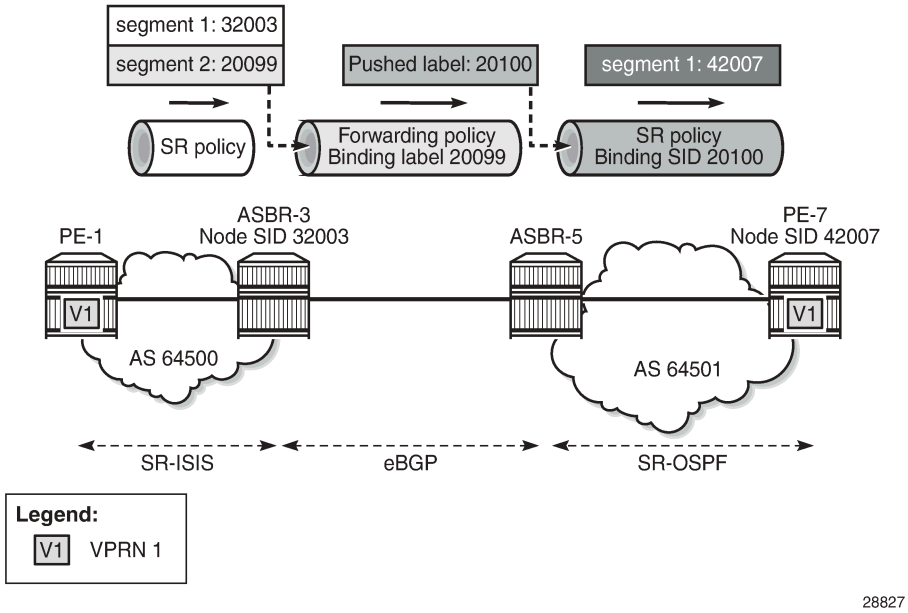
SR policies contain a list of MPLS labels in the form of a segment list that instantiates Segment Routing - Traffic Engineering (SR-TE) Label Switched Paths (LSPs) to a network endpoint and are described in the SR policies chapter.

Inter-AS VPRN Model C using an MPLS forwarding policy and SR policies

One typical application to use MPLS forwarding policies together with SR policies is an example of static Egress Peer Engineering (EPE); a head-end PE in AS1 can steer traffic toward AS2 using a specific AS2 next-hop node.

In the following example, an MPLS forwarding policy is configured on the Autonomous System Border Routers (ASBRs) in an inter-AS VPRN scenario. Figure 9: Inter-AS VPRN Model C using MPLS forwarding policy and SR policies shows the labels added to a packet sent by VPRN 1 on PE-1 in AS 64500 to VPRN 1 on PE-7 in AS 64501.

Figure 9: Inter-AS VPRN Model C using MPLS forwarding policy and SR policies

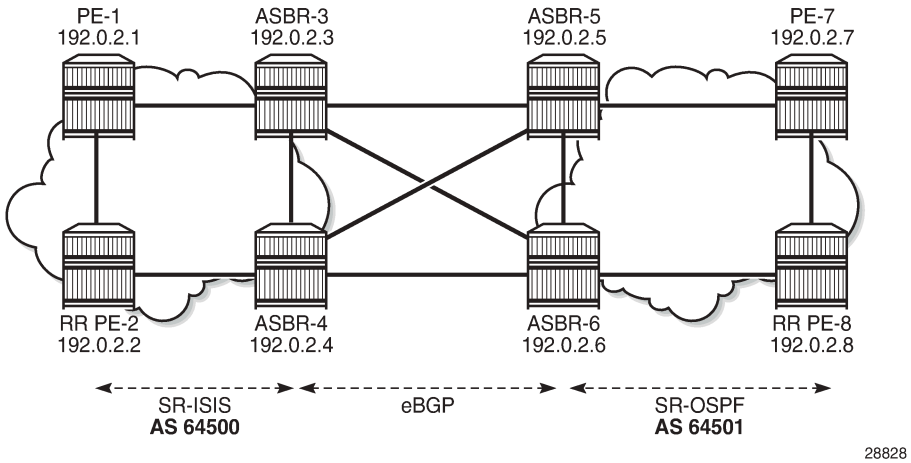


An SR policy on PE-1 pushes two labels: label 32003 from the SRGB for segment routing to SID 32003 of ASBR-3, and label 20099 from the SRLB corresponding to the binding label of the MPLS forwarding policy to be used in ASBR-3. In ASBR-3, these labels are popped and the MPLS forwarding policy is applied. This MPLS forwarding policy forwards the packet to ASBR-5 and pushes a binding label 20100 from the SRLB on ASBR-5, which identifies the SR policy to be used on ASBR-5. In ASBR-5, label 20100 is popped and an SR policy with binding SID 20100 is applied. This SR policy pushes label 42007, which is the SID of PE-7.

Configuration

Figure 10: Example topology shows the example topology with four routers in AS 64500 and four routers in AS 64501. SR-ISIS is configured in AS 64500, while SR-OSPF is configured in AS 64501. The SR policies are configured within the ASs, whereas the MPLS forwarding policies are used to set up tunnels between the ASBRs.

Figure 10: Example topology



The initial configuration includes:

- Cards, MDAs, ports
- Router interfaces
- IS-IS as IGP between the routers in AS 64500; OSPF between the routers in AS 64501

Segment routing

SR-ISIS is configured in AS 64500. The following SR-ISIS configuration on PE-1 shows that the prefix SIDs are taken from the SRGB, which uses labels from 32000 to 32999, and the scope of the router capability advertisement is the area. The system interface has SID index 1, so the node SID label will be start label + index = 32000 + 1 = 32001.

Note:
The SRGB block does not need to have the same start value and end value on each router in the AS, but it must have the same size, that is, the same number of labels in the SRGB.

```
# on PE-1:
configure {
    router "Base" {
        autonomous-system 64500
        mpls-labels {
            sr-labels {
                start 32000
                end 32999   # SRGB block definition AS 64500
            }
        }
        isis 0 {            # IS-IS in the AS 64500; OSPF in ASs 64501
            admin-state enable
            advertise-router-capability area
            level-capability 2
            traffic-engineering true
            area-address [49.0001]
            segment-routing {
                admin-state enable
                prefix-sid-range {
```

```
                    global
                }
            }
            interface "int-PE-1-ASBR-3" {
                interface-type point-to-point
            }
            interface "int-PE-1-PE-2" {
                interface-type point-to-point
            }
            interface "system" {
                ipv4-node-sid {
                    index 1
                }
            }
        }
    }
}
```

The IS-IS configuration is similar on the other routers in AS 64500; the node SID index is different. On the routers in AS 64501, SR-OSPF is configured instead. The SR-OSPF configuration on PE-7 is as follows:

```
# on PE-7:
configure {
    router "Base" {
        autonomous-system 64501
        mpls-labels {
            sr-labels {
                start 42000
                end 42999            # SRGB block definition AS 64501
            }
        }
        ospf 0 {
            admin-state enable
            advertise-router-capability area
            traffic-engineering true
            segment-routing {
                admin-state enable
                prefix-sid-range {
                    global
                }
            }
            area 0.0.0.0 {
                interface "int-PE-7-ASBR-5" {
                    interface-type point-to-point
                }
                interface "int-PE-7-PE-8" {
                    interface-type point-to-point
                }
                interface "system" {
                    node-sid {
                        index 7
                    }
                }
            }
        }
    }
}
```

The OSPF configuration is similar on the other routers in AS 64501; the node SID index is different.

Inter-AS VPRN Model C

The configuration of inter-AS Model C VPRNs is described in the "Inter-AS VPRN Model C" chapter in the *7450 ESS, 7750 SR, 7950 XRS, and VSR Layer 3 Services Advanced Configuration Guide for MD CLI*. PE-2 acts as the Route Reflector (RR) for the "iBGP_grp" group in AS 64500; in AS 64501, PE-8 acts as the RR. Between AS 64500 and AS 64501, eBGP is configured on the ASBRs.

The RR addresses need to be advertised between the ASs. This can be done using IPv4 or labeled IPv4 (with next-hop resolution enabled). No other PE system addresses need to be advertised.

On PE-1, the system IP address 192.0.2.1/32 need not be exported, because no recursive lookup is required. Instead, the VPRN will be configured with auto-bind to the SR policy and the SR policy tunnel can resolve the next-hop of the VPN-IPv4 route. The following configuration shows the BGP configuration for address family VPN-IPv4 on PE-1:

```
# on PE-1:
configure {
    router "Base" {
        bgp {
            split-horizon true
            group "iBGP_grp" {
                type internal
                family {
                    vpn-ipv4 true
                }
            }
            neighbor "192.0.2.2" {
                group "iBGP_grp"
            }
        }
    }
}
```

On the ASBRs, BGP is configured for the labeled IPv4 address family only. On ASBR-3 and ASBR-4, the BGP next-hop for the labeled IPv4 address family can be resolved using SR-ISIS within AS 64500; on ASBR-5 and ASBR-6, the next-hop can be resolved using SR-OSPF within AS 64501. The forwarding between the ASBRs is based on label-binding MPLS forwarding policies. On ASBR-3, BGP is configured as follows:

```
# on ASBR-3:
configure {
    router "Base" {
        bgp {
            split-horizon true
            next-hop-resolution {
                labeled-routes {
                    transport-tunnel {
                        family label-ipv4 {
                            resolution-filter {
                                sr-isis true
                            }
                        }
                    }
                }
            }
            group "eBGP_grp" {
                peer-as 64501
                advertise-inactive true
                family {
```

```
                    label-ipv4 true
                }
                local-as {
                    as-number 64500
                }
            }
            group "iBGP_grp" {
                type internal
                family {
                    label-ipv4 true
                }
            }
            neighbor "192.0.2.2" {
                group "iBGP_grp"
            }
            neighbor "192.168.35.2" {
                group "eBGP_grp"
            }
            neighbor "192.168.36.2" {
                group "eBGP_grp"
            }
        }
    }
}
```

PE-2 is RR in AS 64500 and has a VPN-IPv4 session with PE-1 and labeled IPv4 sessions with both ASBR-3 and ASBR-4. The BGP next-hop for the labeled IPv4 family can be resolved using SR-ISIS. Between the RRs PE-2 and PE-8, a multi-hop eBGP session is established for the VPN-IPv4 address family. The BGP configuration on PE-2 is as follows:

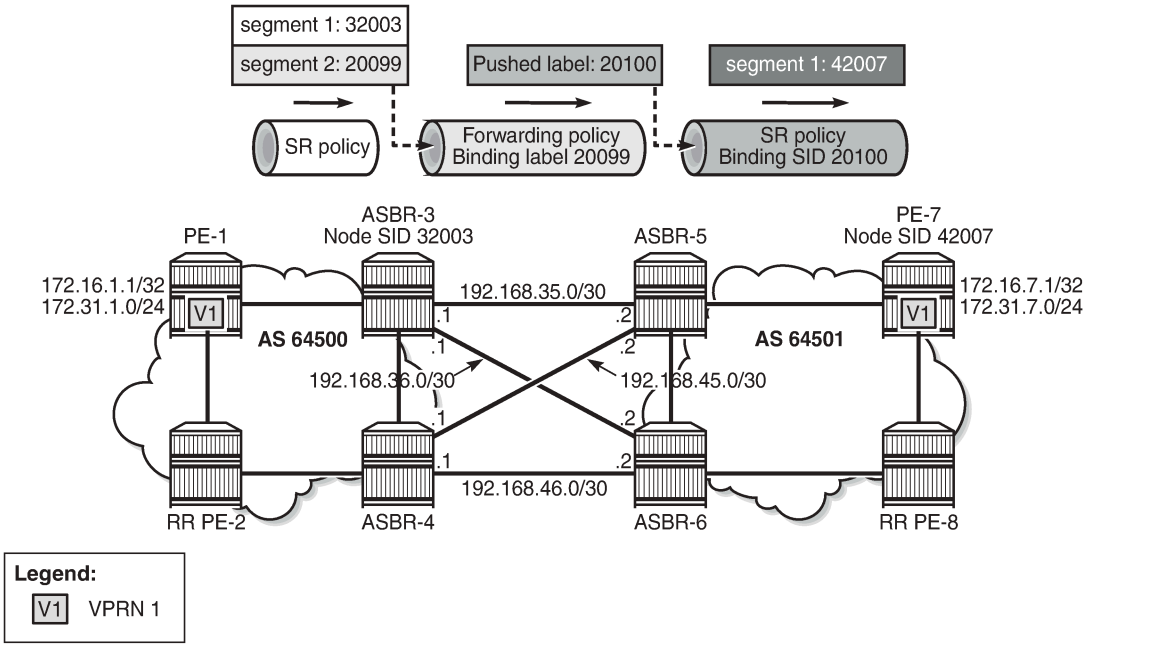
```
# on PE-2:
configure {
    router "Base" {
        bgp {
            split-horizon true
            next-hop-resolution {
                labeled-routes {
                    transport-tunnel {
                        family label-ipv4 {
                            resolution-filter {
                                sr-isis true
                            }
                        }
                    }
                }
            }
            group "eBGP_grp" {
                multihop 10
                type external
                peer-as 64501
                local-address 192.0.2.2
                family {
                    vpn-ipv4 true
                }
                local-as {
                    as-number 64500
                }
            }
            group "iBGP_grp" {
                type internal
                advertise-inactive true
                cluster {
                    cluster-id 192.0.2.2    # PE-2 is RR in AS64500
```

```
                }
            }
            neighbor "192.0.2.1" {
                group "iBGP_grp"
                family {
                    vpn-ipv4 true
                }
            }
            neighbor "192.0.2.3" {
                group "iBGP_grp"
                family {
                    label-ipv4 true
                }
                export {
                    policy ["sysPE_pol"]
                }
            }
            neighbor "192.0.2.4" {
                group "iBGP_grp"
                family {
                    label-ipv4 true
                }
                export {
                    policy ["sysPE_pol"]
                }
            }
            neighbor "192.0.2.8" {            # set up multihop session to remote RR
                group "eBGP_grp"
            }
        }
    }
}
```

For this example, SR policies are used within an AS to resolve VPN-IPv4 next-hops using **auto-bind-tunnel**, whereas MPLS forwarding policies are used between ASs, on the ASBRs. However, SR policies can include a segment with a binding label value that identifies an MPLS forwarding policy defined in one of the local next-hop ASBRs (ASBR-3 or ASBR-4). In addition, MPLS forwarding policies can push a label identifying a BSID in an SR policy defined at the remote next-hop ASBR (ASBR-5 or ASBR-6).

Figure 11: Inter-AS VPRN using MPLS forwarding policy and SR policies: Traffic to PE-7 shows the example topology for an inter-AS VPRN Model C using MPLS forwarding policies and SR policies. For traffic from VPRN 1 on PE-1 to VPRN 1 on PE-7, PE-1 pushes two labels: segment 1 has label 32003, which is the node SID of ASBR-3, and label 20099, which is the binding label of an MPLS forwarding policy defined on ASBR-3. This MPLS forwarding policy on ASBR-3 pops incoming label 20099 and pushes label 20100, which corresponds to the BSID of an SR policy defined on ASBR-5 and ASBR-6. This SR policy pops incoming label 20100 and pushes label 42007, which is the node SID of PE-7.

Figure 11: Inter-AS VPRN using MPLS forwarding policy and SR policies: Traffic to PE-7



SR policies -like MPLS forwarding policies- use labels from an SRLB, which is a pool of labels defined as follows:

```
# on PE-1:
configure {
    router "Base" {
        mpls-labels {
            reserved-label-block "SRLB1" {
                start-label 20000
                end-label 21999
            }
        }
    }
}
```

SR policies contain MPLS labels in a segment list that instantiates SR-TE LSPs to a network endpoint. This appears in the tunnel table as an SR policy tunnel. On PE-1, the following SR policy-with endpoint 192.0.2.7 in a remote AS-is configured with one segment list including two segments:

- segment 1 contains label 32003 referencing the node SID of ASBR-3.
- segment 2 contains label 20099 referencing the binding label that matches the MPLS forwarding policy used at ASBR-3.

There are two ways to steer a set of flows into an SR policy: either based on the BSID value or based on a match of color and endpoint.

In this case, for a VPRN with auto-bind-tunnel, the payload prefix (VPN-IPv4 route prefix) must contain a color community value that matches the color value of the SR-policy route, and the prefix BGP next-hop must also match the endpoint value of the SR-policy route.

In addition, the reserved label block SRLB1 must be referenced within the SR policy context because the configured BSID (20000) is checked against this label block.

```
# on PE-1:
configure {
    router "Base" {
        segment-routing {
            sr-policies {
                admin-state enable
                reserved-label-block "SRLB1"
                static-policy "SR-static-policy-EP7" {
                    admin-state enable
                    color 100
                    endpoint 192.0.2.7
                    head-end local
                    binding-sid 20000
                    distinguisher 64500
                    segment-list 1 {
                        admin-state enable
                        segment 1 {
                            mpls-label 32003    # node SID of ASBR-3
                        }
                        segment 2 {
                            mpls-label 20099    # binding label of fwd-policy
                        }
                    }
                }
            }
        }
    }
}
```

On the ASBRs, MPLS forwarding policies are configured. Like SR policies (using BSID), the binding label is taken from reserved label block SRLB1, which is a pool of labels defined as follows:

```
# on ASBR-3:
configure {
    router "Base" {
        mpls-labels {
            reserved-label-block "SRLB1" {
                start-label 20000
                end-label 21999
            }
        }
    }
}
```

The reserved label block SRLB1 must be referenced within the MPLS forwarding policy context on each ASBR. The following MPLS forwarding policy is configured on ASBR-3 with binding label 20099, which maps to the segment 2 label defined in the SR policy on PE-1. The resolution type is set to direct, meaning that the next-hops are locally attached interface IP addresses. The primary next-hop 192.168.35.2 is on ASBR-5 and the backup next-hop 192.168.36.2 is on ASBR-6.

Within the MPLS forwarding policy on ASBR-3, the pushed label 20100 matches the BSID identifying the SR policy at the peer ASBRs (PE-5 and PE-6). In the MPLS forwarding policy on ASBR3, both the primary and backup next-hops are configured with the same pushed label 20100:

```
# on ASBR-3:
configure {
    router "Base" {
```

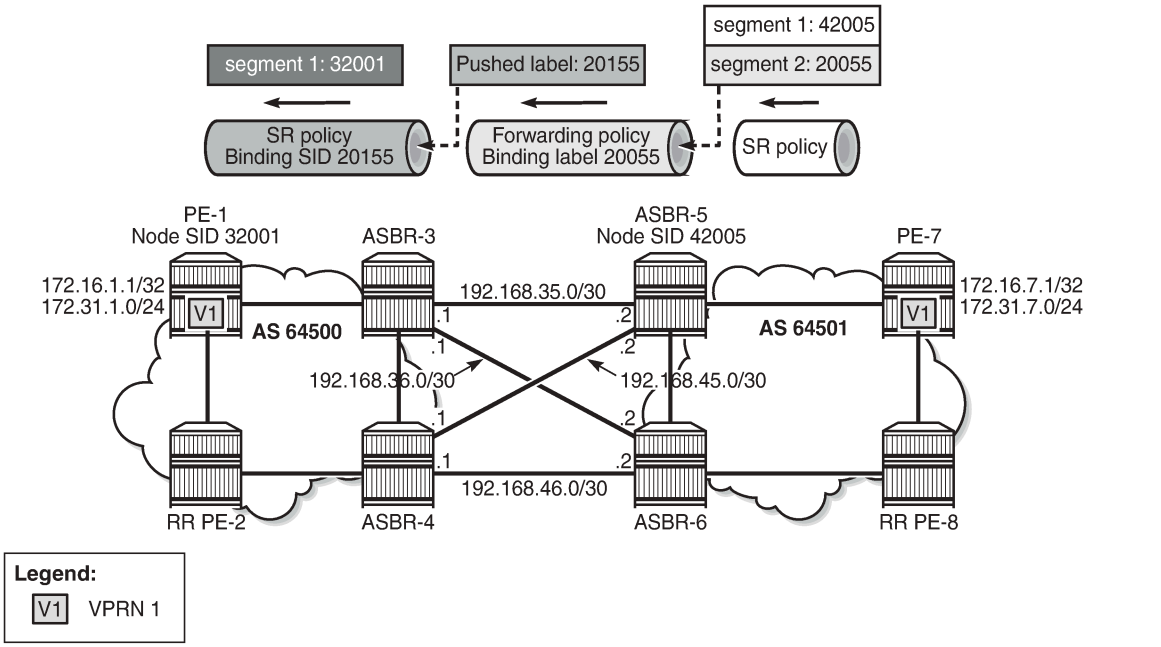
```
        mpls {
            admin-state enable
            forwarding-policies {
                admin-state enable
                reserved-label-block "SRLB1"
                forwarding-policy "SLR-ILM-pushed-label" {
                    admin-state enable
                    binding-label 20099
                    revert-timer 5
                    next-hop-group 1 {
                        admin-state enable
                        resolution-type direct
                        primary-next-hop {
                            next-hop 192.168.35.2
                            pushed-labels 1 {
                                label 20100
                            }
                        }
                        backup-next-hop {
                            next-hop 192.168.36.2
                            pushed-labels 1 {
                                label 20100
                            }
                        }
                    }
                }
            }
        }
    }
}
```

On ASBR-5 and ASBR-6, the following SR policy with endpoint PE-7 and binding SID 20100 only contains one segment toward the node SID of PE-7:

```
# on ASBR-5:
configure {
    router "Base" {
        segment-routing {
            sr-policies {
                admin-state enable
                reserved-label-block "SRLB1"
                static-policy "SR-static-policy-EP7" {
                    admin-state enable
                    color 100
                    endpoint 192.0.2.7
                    head-end local
                    binding-sid 20100
                    distinguisher 64501
                    segment-list 1 {
                        admin-state enable
                        segment 1 {
                            mpls-label 42007    # node SID of PE-7
                        }
                    }
                }
            }
        }
    }
}
```

In the opposite direction, the configuration is similar. Figure 12: Inter-AS VPRN using MPLS forwarding policy and SR policies: Traffic to PE-1 shows the labels used for traffic from PE-7 to PE-1.

Figure 12: Inter-AS VPRN using MPLS forwarding policy and SR policies: Traffic to PE-1



On PE-7, an SR policy is created with endpoint 192.0.2.1 and a segment list with two segments: segment 1 contains label 42005, which is the node SID of ASBR-5, and segment 2 contains label 20055 identifying the binding label in the MPLS forwarding policy at ASBR-5. The configured SR policy color is 100, so this applies for VPN-IPv4 routes with color extended community color:00:100.

```
# on PE-7:
configure {
    router "Base" {
        segment-routing {
            sr-policies {
                admin-state enable
                reserved-label-block "SRLB1"
                static-policy "SR-static-policy-EP1" {
                    admin-state enable
                    color 100
                    endpoint 192.0.2.1
                    head-end local
                    binding-sid 20001
                    distinguisher 64501
                    segment-list 1 {
                        admin-state enable
                        segment 1 {
                            mpls-label 42005    # node SID of ASBR-5
                        }
                        segment 2 {
                            mpls-label 20055    # binding label of fwd-policy
                        }
                    }
                }
            }
        }
    }
```

```
}
```

On ASBR-5, both labels (42005 and 20055) are popped and the MPLS forwarding policy with binding label 20055 pushes label 20155 to the primary and backup next-hops. The configuration is as follows:

```
# on ASBR-5:
configure {
    router "Base" {
        mpls {
            admin-state enable
            forwarding-policies {
                admin-state enable
                reserved-label-block "SRLB1"
                forwarding-policy "SLR-ILM-pushed-label" {
                    admin-state enable
                    binding-label 20055
                    revert-timer 5
                    next-hop-group 1 {
                        admin-state enable
                        resolution-type direct
                        primary-next-hop {
                            next-hop 192.168.35.1
                            pushed-labels 1 {
                                label 20155
                            }
                        }
                        backup-next-hop {
                            next-hop 192.168.45.1
                            pushed-labels 1 {
                                label 20155
                            }
                        }
                    }
                }
            }
        }
    }
}
```

On ASBR-3 and ASBR-4, an SR policy with BSID 20155 is configured. The segment list only contains one segment, which is the node SID of PE-1.

```
# on ASBR-3:
configure {
    router "Base" {
        segment-routing {
            sr-policies {
                admin-state enable
                reserved-label-block "SRLB1"
                static-policy "SR-static-policy-EP1" {
                    admin-state enable
                    color 100
                    endpoint 192.0.2.1
                    head-end local
                    binding-sid 20155
                    distinguisher 64500
                    segment-list 1 {
                        admin-state enable
                        segment 1 {
                            mpls-label 32001    # node SID of PE-1
                        }
                    }
```

```
                }
            }
        }
    }
}
```

On PE-1, VPRN 1 is configured with two loopback interfaces to test the traffic, as follows. The tunnel resolution filter within the service is set to **sr-policy**, configured explicitly. The configuration of VPRN 1 on PE-7 is similar, also with two loopback interfaces for test purposes: 17.16.7.1/32 and 172.31.7.1/24.

```
# on PE-1:
configure {
    service {
        vprn "VPRN 1" {
            admin-state enable
            service-id 1
            customer "1"
            bgp-ipvpn {
                mpls {
                    admin-state enable
                    route-distinguisher "192.0.2.1:1"
                    vrf-target {
                        import-community "target:64501:1"
                        export-community "target:64500:1"
                    }
                    auto-bind-tunnel {
                        resolution filter
                        resolution-filter {
                            sr-policy true
                        }
                    }
                }
            }
            interface "lo1" {
                loopback true
                ipv4 {
                    primary {
                        address 172.31.1.1
                        prefix-length 24
                    }
                }
            }
            interface "system" {
                loopback true
                ipv4 {
                    primary {
                        address 172.16.1.1
                        prefix-length 32
                    }
                }
            }
        }
    }
}
```

The following tunnel table show command on PE-1 returns one tunnel toward PE-7: an SR policy tunnel with color 100.

```
[/]
A:admin@PE-1# show router tunnel-table 192.0.2.7

===============================================================================
```

```
IPv4 Tunnel Table (Router: Base)
===============================================================================
Destination           Owner     Encap TunnelId  Pref   Nexthop        Metric
   Color
-------------------------------------------------------------------------------
192.0.2.7/32          sr-policy MPLS  917506    14     192.0.2.3      0
   100
-------------------------------------------------------------------------------
Flags: B = BGP or MPLS backup hop available
       L = Loop-Free Alternate (LFA) hop available
       E = Inactive best-external BGP route
       k = RIB-API or Forwarding Policy backup hop
===============================================================================
```

The following command shows that the SR policy tunnel with tunnel ID 917506 has next-hop 192.0.2.3 and label 20099, which matches the binding label value of the configured MPLS forwarding policy at next-hop ASBR-3.

```
[/]
A:admin@PE-1# show router fp-tunnel-table 1

===============================================================================
IPv4 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                Protocol            Tunnel-ID
  Lbl/SID
    NextHop                                                    Intf/Tunnel
  Lbl/SID (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
192.0.2.2/32                               SR-ISIS-0            524290
  32002
    192.168.12.2                                               1/1/2:1000
192.0.2.3/32                               SR-ISIS-0            524292
  32003
    192.168.13.2                                               1/1/1:1000
192.0.2.4/32                               SR-ISIS-0            524293
  32004
    192.168.12.2                                               1/1/2:1000
192.0.2.7/32                               SR-Policy            917506
  20099
    192.0.2.3                                                  SR
192.168.12.2/32                            SR                   524289
  3
    192.168.12.2                                               1/1/2:1000
192.168.13.2/32                            SR                   524291
  3
    192.168.13.2                                               1/1/1:1000
-------------------------------------------------------------------------------
Total Entries : 6
-------------------------------------------------------------------------------
===============================================================================
```

However, on PE-1, the received BGP-VPN routes are not used, as follows:

```
[/]
A:admin@PE-1# show router bgp routes vpn-ipv4
===============================================================================
```

```
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================
 Legend -
 Status codes  : u - used, s - suppressed, h - history, d - decayed, * - valid
                 l - leaked, x - stale, > - best, b - backup, p - purge
 Origin codes  : i - IGP, e - EGP, ? - incomplete

===============================================================================
BGP VPN-IPv4 Routes
===============================================================================
Flag  Network                                          LocalPref   MED
      Nexthop (Router)                                 Path-Id     IGP Cost
      As-Path                                                      Label
-------------------------------------------------------------------------------
*>i   192.0.2.7:1:172.16.7.1/32                        100         None
      192.0.2.7                                        None        0
      64501                                                        524286
*>i   192.0.2.7:1:172.31.7.0/24                        100         None
      192.0.2.7                                        None        0
      64501                                                        524286
-------------------------------------------------------------------------------
Routes : 2
===============================================================================
```

Therefore, the following route table for VPRN 1 does not include any route toward VPRN 1 on PE-7:

```
[/]
A:admin@PE-1# show router 1 route-table

===============================================================================
Route Table (Service: 1)
===============================================================================
Dest Prefix[Flags]                             Type    Proto     Age        Pref
      Next Hop[Interface Name]                                     Metric
-------------------------------------------------------------------------------
172.16.1.1/32                                  Local   Local     00h46m33s  0
       system                                                        0
172.31.1.0/24                                  Local   Local     00h46m33s  0
       lo1                                                           0
-------------------------------------------------------------------------------
No. of Routes: 2
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
===============================================================================
```

As stated earlier, the following two conditions are required to ensure that the traffic flow from VPRN 1 is steered by the SR policy on PE-1:

- the BGP payload prefix next-hop must match the endpoint value in the SR policy
- the BGP payload prefix must have a color extended community, matching the color value in the SR policy

To match the second condition, the following BGP policy exports the prefixes of VPRN 1 and adds color extended community "color:00:100" and extended community "target:64500:1" to the VPN-IPv4 routes.

```
# on PE-1:
configure {
    policy-options {
        community "Color100_com" {
            member "color:00:100" { }
```

```
        }
        community "VRF1-export_com" {
            member "target:64500:1" { }
        }
        prefix-list "VRF1-prefixes" {
            prefix 172.16.1.1/32 type exact {
            }
            prefix 172.31.1.0/24 type exact {
            }
        }
        policy-statement "Color100_VRF1_pol" {
            entry 10 {
                from {
                    prefix-list ["VRF1-prefixes"]
                }
                to {
                    protocol {
                        name [bgp-vpn]
                    }
                }
                action {
                    action-type accept
                    community {
                        add ["VRF1-export_com" "Color100_com"]
                    }
                }
            }
        }
    }
}
```

This policy is applied as a VRF export policy in VPRN 1, as follows:

```
# on PE-1:
configure {
    service {
        vprn "VPRN 1" {
            bgp-ipvpn {
                mpls {
                    vrf-export {
                        policy ["Color100_VRF1_pol"]
                    }
                }
            }
        }
    }
}
```

The BGP update message shows that the VPN-IPv4 routes contain both extended communities:

```
1 2021/09/20 14:31:41.065 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.2
"Peer 1: 192.0.2.2: UPDATE
Peer 1: 192.0.2.2 - Send BGP UPDATE:
    Withdrawn Length = 0
    Total Path Attr Length = 85
    Flag: 0x90 Type: 14 Len: 48 Multiprotocol Reachable NLRI:
        Address Family VPN_IPV4
        NextHop len 12 NextHop 192.0.2.1
        172.16.1.1/32 RD 192.0.2.1:1 Label 524287 (Raw label 0x7ffff1)
        172.31.1.0/24 RD 192.0.2.1:1 Label 524287 (Raw label 0x7ffff1)
    Flag: 0x40 Type: 1 Len: 1 Origin: 0
    Flag: 0x40 Type: 2 Len: 0 AS Path:
    Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
```

```
    Flag: 0xc0 Type: 16 Len: 16 Extended Community:
        target:64500:1
        color:00:100
"
```

The VPN-IPv4 routes are using the SR policy tunnel to PE-7, as follows:

```
[/]
A:admin@PE-1# show router bgp next-hop vpn-ipv4
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================

===============================================================================
BGP VPN Next Hop
===============================================================================
VPN Next Hop                                                  Owner
   Autobind                                          FibProg  Reason
   Labels                                            FlexAlgo Metric
   Admin-tag-policy (strict-tunnel-tagging)
-------------------------------------------------------------------------------
192.0.2.7                                                     SR-POLICY
   bgp sr-policy                                     Y
   --                                                --       0
   -- (-)
-------------------------------------------------------------------------------
Next Hops : 1
===============================================================================
```

In the route table of VPRN 1 on PE-1, routes to VPRN 1 on PE-7 use the SR policy tunnel toward PE-7, as follows:

```
[/]
A:admin@PE-1# show router 1 route-table

===============================================================================
Route Table (Service: 1)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
172.16.1.1/32                                 Local   Local     01h12m28s  0
       system                                                        0
172.16.7.1/32                                 Remote  BGP VPN   00h17m00s  170
       192.0.2.7 (tunneled:SR-Policy:917506)                         0
172.31.1.0/24                                 Local   Local     01h12m28s  0
       lo1                                                           0
172.31.7.0/24                                 Remote  BGP VPN   00h17m00s  170
       192.0.2.7 (tunneled:SR-Policy:917506)                         0
-------------------------------------------------------------------------------
No. of Routes: 4
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
===============================================================================
```

Note:
Color-only steering can be achieved without the need to match the BGP next-hop endpoint. This is done by setting the "color-only" (CO) bits, which are the two highest order bits of the color extended community. When set to "10" or "01" instead of "00", the **endpoint** value in the SR

policy can be set to "0.0.0.0". In this case, only the color value is checked as a single condition to steer traffic flows using the SR policy.

Show and Debug commands

The following command shows the SR policies on PE-1, with the segment list, color, endpoint address, and so on:

```
[/]
A:admin@PE-1# show router segment-routing sr-policies all

===============================================================================
SR-Policies Path
===============================================================================
-------------------------------------------------------------------------------
Active          : Yes                   Owner           : static
Color           : 100
Head            : 0.0.0.0               Endpoint Addr   : 192.0.2.7
RD              : 64500                 Preference      : 100
BSID            : 20000
TunnelId        : 917506                Age             : 218
Origin ASN      : 0                     Origin          : 0.0.0.0
NumReEval       : 0                     ReEvalReason    : none
NumActPathChange: 0                     Last Change     : 09/20/2021 14:18:28
Maintenance Policy: N/A

Path Segment Lists:
Segment-List    : 1                     Weight          : 1
S-BFD State     : Down                  S-BFD Transitio*: 0
Num Segments    : 2                     Last Change     : 09/20/2021 13:31:51
  Seg 1 Label   : 32003                 State           : resolved-up
  Seg 2 Label   : 20099                 State           : N/A

===============================================================================
* indicates that the corresponding row element may have been truncated.
```

For each combination of color and endpoint, the SR database must validate each segment list / candidate path, and choose one to be the active path. The most important checks are:

- The configured BSID is part of the SRLB. In this example, 20000 is part of the SRLB1 block on PE-1, which ranges from 20000 to 21999. The BSID is used uniquely by this policy.
- The first segment of each segment list is resolved to a set of one or more next-hops. This means matching an SR-ISIS or SR-OSPF node SID, matching an SR-ISIS or SR-OSPF adjacency SID, or matching an SR-ISIS or SR-OSPF adjacency-set SID. In this example, on PE-1, label 32003 resolves to the SR-ISIS node SID of ASBR-3.

The following command shows the MPLS forwarding policy on ASBR-3, including the binding label, NHG, and pushed labels. If, for example, the reserved label block was not defined or not referenced by the MPLS forwarding policy, the validation would fail and the MPLS forwarding policy would remain operationally down.

```
[/]
A:admin@ASBR-3# show router mpls forwarding-policies forwarding-policy detail

===============================================================================
Forwarding Policy (Detail)
===============================================================================
-------------------------------------------------------------------------------
```

```
Policy : SLR-ILM-pushed-label
------------------------------------------------------------------------------
Admin State      : Up                   Oper State       : Up
Binding Label    : 20099                Preference       : 255
Revert Timer     : 5 sec
Last Change      : 09/20/2021 14:19:00
Ingress Stats    : Disabled
Metric           : 0                    Tunnel Table Pref: 255
Endpoint Address : N/A

Next-hop Group   : 1
Admin State      : Up                   Oper State       : Up
Resolution Type  : direct               Load Balancing Wt: 0
Last Change      : 09/20/2021 14:19:00
Primary
NH Address       : 192.168.35.2
Oper State       : Up                   Last Change      : 09/20/2021 14:19:00
Pushed Labels    : 20100
Backup
NH Address       : 192.168.36.2
Oper State       : Up                   Last Change      : 09/20/2021 14:19:00
Pushed Labels    : 20100
==============================================================================
```

The following command shows the details of the MPLS binding label forwarding policy:

```
[/]
A:admin@ASBR-3# show router mpls forwarding-policies binding-label detail

==============================================================================
Binding Label (Detail)
==============================================================================
Label          : 20099                  Preference     : 255
Policy Name    : SLR-ILM-pushed-label
Oper State     : Up                     OperDownReason : notApplicable
Up Time        : 09/20/2021 14:07:03    NumNextHopGrps : 1
Ingress Stats  : Disabled               IngrOperState  : Down
Egress Stats   : Disabled               EgrOperState   : Down
Revert Timer   : 5
Retry Count    : 0                      Next Retry In  : 0

Next-hop Group : 1                      Resolution Type: direct
Oper State     : Up                     OperDownReason : notApplicable
Num Revert     : 0                      Num Failover   : 0
Next Revert In : 0
Primary nexthop: 192.168.35.2
Resolved       : True                   NHopDownReason : notApplicable
EgrOperState   : Down
Pushed Labels  : 20100
Backup nexthop : 192.168.36.2
Resolved       : True                   NHopDownReason : notApplicable
EgrOperState   : Down
Pushed Labels  : 20100
------------------------------------------------------------------------------
==============================================================================
```

The following tools command on ASBR-3 shows that the MPLS forwarding policy is validated and the ILM is programmed with the binding label value. The output also shows which router interfaces are used toward the configured next-hops and what label stack is pushed.

```
[/]
A:admin@ASBR-3# tools dump router mpls forwarding-policies binding-label 20099
```

```
 Db Mgr flags 0x80 ilmStatsFailCnt 0
-------------------------------------------------------------------------------
 dbOwner FWD PLCY routeOwner 48 rsvdBlkId 2 flags 0x3 numPolicies 1 numInstalled 1
-------------------------------------------------------------------------------

 Label DB 20099
 dbFlags 0xd PathCount 1 srTunnelId 851970 ilmStatsIdx[MGMT] 0x0 ilmStatsIdx[API] 0x0
LABEL RESERVED:  PROGRAMMED
 Path bitmap 0
 Label Retry time left : 0 retrycount : 0,  SR Retry time left : 0 SR retrycount : 0

 Best Db Path owner 0 path name vrId:1, dbOwner:0, pathName:SLR-ILM-pushed-label Last Modified
 09/20/2021 14:19:00 Up Time 0d 00:33:31
 Preference 255 flags 0x245 Status FWDPLCY_ERR_NA SR status SR_ERR_OK
 PrimResolved NH's 1 BkupResolved NH's 1
 NHGroup 1
 flags 0x3bf9 : weight 0 normalized weight 0
 Revert timer 5 Time left 0 NumOfReverts 0

 Hold timer 0 Time left 0
DIRECT NH:  PRIM PGMED:  PRIM RESOLVED:  BKUP RESOLVED:  BKUP PGMED:
 primaryNH 192.168.35.2 egrStatsIdx 0x0 Status FWDPLCY_NHERR_NA
 Label Stack:20100 0
 Nexthop 1 192.168.35.2 outIf 3 globalIfIndex 2 globaIfInNHgrp 2
 PG ID 4
 PG ID 0
 backupNH 192.168.36.2 egrStatsIdx 0x0 Status FWDPLCY_NHERR_NA
 Label stack:20100 0
 Nexthop 1 192.168.36.2 outIf 4 globalIfIndex 3 globaIfInNHgrp 3
 PG ID 5
 PG ID 0

 -------------------------------------------------------------------------
```

Conclusion

MPLS forwarding policies provide the customization of next-hops as well as ECMP, weighted ECMP, Class-Based Forwarding, and backup support. MPLS forwarding policies can be combined with SR policies.

Parallel Adjacency Sets in Segment Routing

This chapter describes the Parallel Adjacency Sets in Segment Routing.

Topics in this chapter include:

Applicability

The information and configuration in this chapter are based on SR OS Release 21.7.R1. They apply for MD-CLI.

Overview

SR OS supports segment routing as described in RFC 8402, *Segment Routing Architecture*. In the remainder of this chapter, SR refers to "Segment Routing", unless specified otherwise. Product and release references, such as 7750 SR and SR OS, continue to refer to "Service Router".

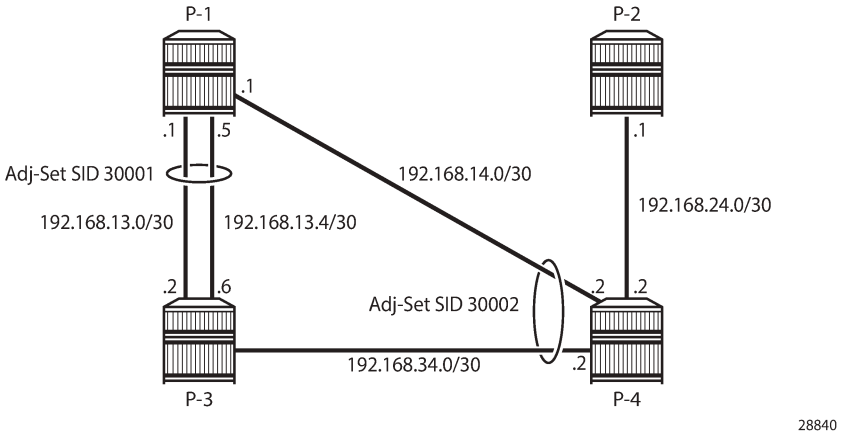
SR provides operators the means to provision paths or tunnels, encoded as a sequential list of sub-paths or segments without requiring a dedicated signaling protocol, by advertising the identities of the segments across the SR domain using extensions to the link state Interior Gateway Protocols (IGPs), such as IS-IS and OSPF.

When defining source-routed traffic-engineered end-to-end SR paths, routing constraints such as loose and strict hops can be used to control the data path through a network; a node SID is used for a loose hop, and an adjacency SID is used for a strict hop. See the Segment Routing – Traffic Engineered Tunnels chapter for more information.

Parallel links between adjacent nodes can be grouped into adjacency sets, and a single adjacency set is identified using a locally significant adjacency set SID. Traffic can be load shared across the links in the set and is based on traffic flow identifiers; for example, source and destination IP addresses, and entropy label.

In Figure 13: Parallel and non-parallel adjacency sets, two adjacency sets are defined. A first adjacency set is defined on P-1 with adjacency set SID 30001. Two parallel links are available between P-1 and P-2, and by combining them into an adjacency set, traffic can be shared across both links. A second set is defined on P-4, with adjacency set SID 30002. However, the member links of that set are not terminated on the same router pair, so traffic cannot be shared.

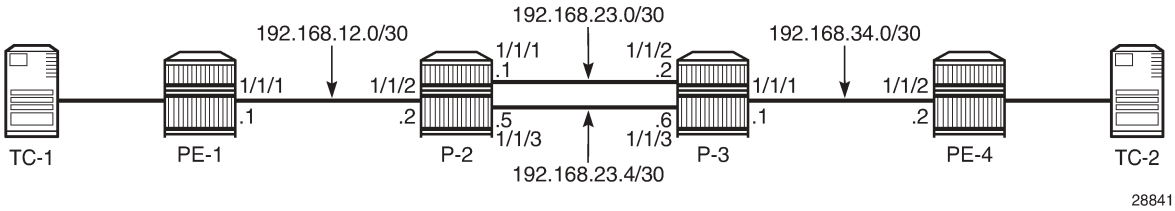
Figure 13: Parallel and non-parallel adjacency sets



Configuration

The topology used in this chapter is shown in Figure 14: Parallel adjacency set. All nodes are configured for SR and IS-IS level 2. If test center TC-1 is connected at PE-1 and test center TC-2 is connected at PE-4, traffic can be sent from TC-1 to TC-2 following the PE-1, P-2, P-3, PE-4 path. Two links are active between P-2 and P-3, and these links belong to the same adjacency set.

Figure 14: Parallel adjacency set



The initial configuration on the PE nodes includes the following:

- Cards, MDAs, ports
- Router interfaces
- IS-IS

Segment routing configuration

In the topology from Figure 14: Parallel adjacency set, all nodes are configured with a common Segment Routing Global Block (SRGB), which is defined as follows:

```
configure {
    router "Base" {
        mpls-labels {
            sr-labels {
```

```
                start 20000
                end 20099
            }
        }
    }
}
```

In this example, prefix SID allocation is using global mode, and the node SIDs are defined by index on the system interfaces in the **isis** context, where PE-1, P-2, P-3, and PE-4 take the indices 1, 2, 3, and 4, respectively. The **advertise-router-capability area** command enables the IS-IS extensions so that the SID values are advertised throughout the SR domain. The configuration on PE-1 is as follows; the configuration on the other nodes is similar.

```
# on PE-1
configure {
    router "Base" {
        isis 0 {
            admin-state enable
            advertise-router-capability area
            level-capability 2
            traffic-engineering true
            area-address [49.0001]
            segment-routing {
                admin-state enable
                prefix-sid-range {
                    global
                }
            }
            interface "int-PE-1-P-2" {
                interface-type point-to-point
            }
            interface "system" {
                ipv4-node-sid {
                    index 1
                }
            }
        }
    }
}
```

With this configuration, each node floods the SIDs in link state packets (shown as "LSP") across the domain. For P-2, prefix 192.0.2.2 has index 2 in the SRGB. The adjacency SIDs 524285, 524286, and 524287 are taken from the dynamic range, as follows:

```
[/]
A:admin@P-2# show router isis database P-2.00-00 detail

===============================================================================
Rtr Base ISIS Instance 0 Database (detail)
===============================================================================

Displaying Level 1 database
-------------------------------------------------------------------------------
Level (1) LSP Count : 0

Displaying Level 2 database
-------------------------------------------------------------------------------
LSP ID    : P-2.00-00                                 Level     : L2
Sequence  : 0x3                   Checksum  : 0xeb4b   Lifetime  : 1103
Version   : 1                     Pkt Type  : 20       Pkt Ver   : 1
Attributes: L1L2                  Max Area  : 3        Alloc Len : 1492
```

```
SYS ID    : 1920.0000.2002         SysID Len : 6        Used Len  : 330

TLVs :
  Area Addresses:
    Area Address : (3) 49.0001
  Supp Protocols:
    Protocols     : IPv4
  IS-Hostname   : P-2
  Router ID   :
    Router ID   : 192.0.2.2
  Router Cap : 192.0.2.2, D:0, S:0
    TE Node Cap : B E M  P
    SR Cap: IPv4 MPLS-IPv6
       SRGB Base:20000, Range:100
    SR Alg: metric based SPF
    Node MSD Cap: BMI : 12 ERLD : 15
  IS Neighbors  :
    Virtual Flag  : 0
    Default Metric: (I) 10
    Delay Metric  : (I) 0
    Expense Metric: (I) 0
    Error Metric  : (I) 0
    Neighbor      : PE-1.00
  IS Neighbors  :
    Virtual Flag  : 0
    Default Metric: (I) 10
    Delay Metric  : (I) 0
    Expense Metric: (I) 0
    Error Metric  : (I) 0
    Neighbor      : P-3.00
  Internal Reach:
    Default Metric: (I) 10
    Delay Metric  : (I) 0
    Expense Metric: (I) 0
    Error Metric  : (I) 0
    IP Address    : 192.168.12.0
    IP Mask       : 255.255.255.252
    Default Metric: (I) 10
    Delay Metric  : (I) 0
    Expense Metric: (I) 0
    Error Metric  : (I) 0
    IP Address    : 192.168.23.0
    IP Mask       : 255.255.255.252
    Default Metric: (I) 0
    Delay Metric  : (I) 0
    Expense Metric: (I) 0
    Error Metric  : (I) 0
    IP Address    : 192.0.2.2
    IP Mask       : 255.255.255.255
    Default Metric: (I) 10
    Delay Metric  : (I) 0
    Expense Metric: (I) 0
    Error Metric  : (I) 0
    IP Address    : 192.168.23.4
    IP Mask       : 255.255.255.252
  I/F Addresses :
    I/F Address   : 192.168.23.1
    I/F Address   : 192.0.2.2
    I/F Address   : 192.168.12.2
    I/F Address   : 192.168.23.5
  TE IS Nbrs   :
    Nbr   : PE-1.00
    Default Metric  : 10
    Sub TLV Len     : 19
```

```
    IF Addr   : 192.168.12.2
    Nbr IP    : 192.168.12.1
    Adj-SID: Flags:v4VL Weight:0 Label:524287
  TE IS Nbrs   :
    Nbr   : P-3.00
    Default Metric  : 10
    Sub TLV Len     : 26
    IF Addr   : 192.168.23.1
    Nbr IP    : 192.168.23.2
    Adj-SID: Flags:v4VL Weight:0 Label:524286
    Adj-SID: Flags:v4VLSP Weight:0 Label:30000
  TE IS Nbrs   :
    Nbr   : P-3.00
    Default Metric  : 10
    Sub TLV Len     : 26
    IF Addr   : 192.168.23.5
    Nbr IP    : 192.168.23.6
    Adj-SID: Flags:v4VL Weight:0 Label:524285
    Adj-SID: Flags:v4VLSP Weight:0 Label:30000
  TE IP Reach   :
    Default Metric  : 10
    Control Info:    , prefLen 30
    Prefix   : 192.168.12.0
    Default Metric  : 10
    Control Info:    , prefLen 30
    Prefix   : 192.168.23.0
    Default Metric  : 0
    Control Info:   S, prefLen 32
    Prefix   : 192.0.2.2
    Sub TLV   :
      Prefix-SID Index:2, Algo:0, Flags:NnP
    Default Metric  : 10
    Control Info:    , prefLen 30
    Prefix   : 192.168.23.4

Level (2) LSP Count : 1
-------------------------------------------------------------------------------
Control Info     : D = Prefix Leaked Down
                   S = Sub-TLVs Present
Attribute Flags  : N = Node Flag
                   R = Re-advertisement Flag
                   X = External Prefix Flag
                   E = Entropy Label Capability (ELC) Flag
Adj-SID Flags    : v4/v6 = IPv4 or IPv6 Address-Family
                   B = Backup Flag
                   V = Adj-SID carries a value
                   L = value/index has local significance
                   S = Set of Adjacencies
                   P = Persistently allocated
Prefix-SID Flags : R = Re-advertisement Flag
                   N = Node-SID Flag
                   nP = no penultimate hop POP
                   E = Explicit-Null Flag
                   V = Prefix-SID carries a value
                   L = value/index has local significance
Lbl-Binding Flags: v4/v6 = IPv4 or IPv6 Address-Family
                   M = Mirror Context Flag
                   S = SID/Label Binding flooding
                   D = Prefix Leaked Down
                   A = Attached Flag
SABM-flags Flags:  R = RSVP-TE
                   S = SR-TE
                   F = LFA
                   X = FLEX-ALGO
```

```
FAD-flags Flags:   M = Prefix Metric
===============================================================================
```

Adjacency set configuration

Adjacency set SIDs are allocated from a reserved label block. Because the adjacency SIDs have a local significance only, the same block can be defined on each node. In this example, a different label block is defined on P-2 and P-3 respectively, as follows. The start-label and end-label values must be in the dynamic range.

```
# on P-2
configure {
    router "Base" {
        mpls-labels {
            reserved-label-block "adjset_block_on_P-2" {
                start-label 30000
                end-label 30099
            }
        }
    }
}

# on P-3
configure {
    router "Base" {
        mpls-labels {
            reserved-label-block "adjset_block_on_P-3" {
                start-label 40000
                end-label 40099
            }
        }
    }
}
```

This range is listed in the **show router mpls-labels label-range** command, as follows:

```
[/]
A:admin@P-2# show router mpls-labels label-range

===============================================================================
Label Ranges
===============================================================================
Label Type      Start Label End Label   Aging       Available   Total
-------------------------------------------------------------------------------
Static          32          18431       -           18400       18400
Dynamic         18432       524287      0           505653      505856
    Seg-Route   20000       20099       -           0           100

-------------------------------------------------------------------------------
Reserved Label Blocks
-------------------------------------------------------------------------------
Reserved Label                          Start       End         Total
Block Name                              Label       Label
-------------------------------------------------------------------------------
adjset_block_on_P-2                     30000       30099       100
-------------------------------------------------------------------------------
No. of Reserved Label Blocks: 1
-------------------------------------------------------------------------------
===============================================================================
```

The reserved label block range is then defined as a Segment Routing Local Block (SRLB) in the segment-routing context. Label values for adjacency sets must be allocated from the SRLB; otherwise, an error is raised. The adjacency set is identified by number, and on P-2 adjacency set 1 has a SID label value of 30000. A similar configuration is used on P-3. If no SID label value is configured, the system will allocate a value from the SRLB range.

```
# on P-2
configure {
    router "Base" {
        isis 0 {
            segment-routing {
                admin-state enable
                srlb "adjset_block_on_P-2"
                adjacency-set 1 {
                    sid {
                        label 30000
                    }
                }
            }
        }
    }
}

# on P-3
configure {
    router "Base" {
        isis 0 {
            segment-routing {
                admin-state enable
                srlb "adjset_block_on_P-3"
                adjacency-set 1 {
                    sid {
                        label 40000
                    }
                }
            }
        }
    }
}
```

On P-2, the *int-P-2-P-3-a* and *int-P-2-P-3-b* interfaces have addresses 192.168.23.1/30 and 192.168.23.5/30, respectively, and these interfaces are included in adjacency set 1 by applying the adjacency set index to the individual interfaces, as follows. A similar configuration is present on P-3.

```
configure {
    router "Base" {
        isis 0 {
            interface "int-P-2-P-3-a" {
                interface-type point-to-point
                adjacency-set 1 { }
            }
            interface "int-P-2-P-3-b" {
                interface-type point-to-point
                adjacency-set 1 { }
            }
            interface "int-P-2-PE-1" {
                interface-type point-to-point
            }
            interface "system" {
                ipv4-node-sid {
                    index 2
```

```
                }
            }
        }
    }
}
```

With this configuration applied, IS-IS floods the adjacency set SID in the adjacency SID sub-TLV across the domain, as follows. The S flag indicates that this SID identifies an adjacency set; the P flag indicates that the SID value is persistent.

```
[/]
A:admin@P-2# show router isis database P-2.00-00 detail

===============================================================================
Rtr Base ISIS Instance 0 Database (detail)
===============================================================================

Displaying Level 1 database
-------------------------------------------------------------------------------
Level (1) LSP Count : 0

Displaying Level 2 database
-------------------------------------------------------------------------------
LSP ID    : P-2.00-00                                 Level     : L2
Sequence  : 0x3                   Checksum  : 0xeb4b   Lifetime  : 1103
Version   : 1                     Pkt Type  : 20       Pkt Ver   : 1
Attributes: L1L2                  Max Area  : 3        Alloc Len : 1492
SYS ID    : 1920.0000.2002        SysID Len : 6        Used Len  : 330

TLVs :
  Area Addresses:
    Area Address : (3) 49.0001
  Supp Protocols:
    Protocols     : IPv4
  IS-Hostname   : P-2
  Router ID   :
    Router ID   : 192.0.2.2
  Router Cap : 192.0.2.2, D:0, S:0
    TE Node Cap : B E M  P
    SR Cap: IPv4 MPLS-IPv6
       SRGB Base:20000, Range:100
    SR Alg: metric based SPF
    Node MSD Cap: BMI : 12 ERLD : 15
  IS Neighbors  :
    Virtual Flag  : 0
    Default Metric: (I) 10
    Delay Metric  : (I) 0
    Expense Metric: (I) 0
    Error Metric  : (I) 0
    Neighbor      : PE-1.00
  IS Neighbors  :
    Virtual Flag  : 0
    Default Metric: (I) 10
    Delay Metric  : (I) 0
    Expense Metric: (I) 0
    Error Metric  : (I) 0
    Neighbor      : P-3.00
  Internal Reach:
    Default Metric: (I) 10
    Delay Metric  : (I) 0
    Expense Metric: (I) 0
    Error Metric  : (I) 0
    IP Address    : 192.168.12.0
```

```
    IP Mask       : 255.255.255.252
    Default Metric: (I) 10
    Delay Metric  : (I) 0
    Expense Metric: (I) 0
    Error Metric  : (I) 0
    IP Address    : 192.168.23.0
    IP Mask       : 255.255.255.252
    Default Metric: (I) 0
    Delay Metric  : (I) 0
    Expense Metric: (I) 0
    Error Metric  : (I) 0
    IP Address    : 192.0.2.2
    IP Mask       : 255.255.255.255
    Default Metric: (I) 10
    Delay Metric  : (I) 0
    Expense Metric: (I) 0
    Error Metric  : (I) 0
    IP Address    : 192.168.23.4
    IP Mask       : 255.255.255.252
  I/F Addresses :
    I/F Address   : 192.168.23.1
    I/F Address   : 192.0.2.2
    I/F Address   : 192.168.12.2
    I/F Address   : 192.168.23.5
  TE IS Nbrs   :
    Nbr   : PE-1.00
    Default Metric  : 10
    Sub TLV Len     : 19
    IF Addr   : 192.168.12.2
    Nbr IP    : 192.168.12.1
    Adj-SID: Flags:v4VL Weight:0 Label:524287
  TE IS Nbrs   :
    Nbr   : P-3.00
    Default Metric  : 10
    Sub TLV Len     : 26
    IF Addr   : 192.168.23.1
    Nbr IP    : 192.168.23.2
    Adj-SID: Flags:v4VL Weight:0 Label:524286
    Adj-SID: Flags:v4VLSP Weight:0 Label:30000
  TE IS Nbrs   :
    Nbr   : P-3.00
    Default Metric  : 10
    Sub TLV Len     : 26
    IF Addr   : 192.168.23.5
    Nbr IP    : 192.168.23.6
    Adj-SID: Flags:v4VL Weight:0 Label:524285
    Adj-SID: Flags:v4VLSP Weight:0 Label:30000
  TE IP Reach   :
    Default Metric  : 10
    Control Info:    , prefLen 30
    Prefix   : 192.168.12.0
    Default Metric  : 10
    Control Info:    , prefLen 30
    Prefix   : 192.168.23.0
    Default Metric  : 0
    Control Info:   S, prefLen 32
    Prefix   : 192.0.2.2
    Sub TLV   :
      Prefix-SID Index:2, Algo:0, Flags:NnP
    Default Metric  : 10
    Control Info:    , prefLen 30
    Prefix   : 192.168.23.4

Level (2) LSP Count : 1
```

```
-------------------------------------------------------------------------------
Control Info     : D = Prefix Leaked Down
                   S = Sub-TLVs Present
Attribute Flags  : N = Node Flag
                   R = Re-advertisement Flag
                   X = External Prefix Flag
                   E = Entropy Label Capability (ELC) Flag
Adj-SID Flags    : v4/v6 = IPv4 or IPv6 Address-Family
                   B = Backup Flag
                   V = Adj-SID carries a value
                   L = value/index has local significance
                   S = Set of Adjacencies
                   P = Persistently allocated
Prefix-SID Flags : R = Re-advertisement Flag
                   N = Node-SID Flag
                   nP = no penultimate hop POP
                   E = Explicit-Null Flag
                   V = Prefix-SID carries a value
                   L = value/index has local significance
Lbl-Binding Flags: v4/v6 = IPv4 or IPv6 Address-Family
                   M = Mirror Context Flag
                   S = SID/Label Binding flooding
                   D = Prefix Leaked Down
                   A = Attached Flag
SABM-flags Flags:  R = RSVP-TE
                   S = SR-TE
                   F = LFA
                   X = FLEX-ALGO
FAD-flags Flags:   M = Prefix Metric
===============================================================================
```

SR traffic engineered label switched path configuration

For traffic from PE-1 to PE-4 to use the adjacency set between P-2 and P-3, a label switched path is required. This path can be defined using SR policies or using SR traffic engineered (SR-TE) tunnels (see the Segment Routing – Traffic Engineered Tunnels chapter).

This chapter uses SR-TE tunnels, with label switched path *lsp-adj-set* using *path-adj-set* as the primary path. A loose hop translates to a node SID for that hop. A strict hop translates to an adjacency set SID, if an adjacency set is available. If no adjacency set is configured, an adjacency SID is used. The MPLS configuration on PE-1 is as follows; the configuration on PE-4 is similar.

```
# on PE-1
configure {
    router "Base" {
        mpls {
            admin-state enable
            path "path-adj-set" {
                admin-state enable
                hop 1 {
                    ip-address 192.0.2.2
                    type loose
                }
                hop 2 {
                    ip-address 192.0.2.3
                    type strict
                }
                hop 3 {
                    ip-address 192.0.2.4
                    type loose
```

```
                }
            }
            lsp "lsp-adj-set" {
                admin-state enable
                type p2p-sr-te
                to 192.0.2.4
                max-sr-labels {
                    label-stack-size 3
                    additional-frr-labels 2
                }
                primary "path-adj-set" {
                }
            }
        }
    }
}
```

The path details for the *lsp-adj-set* SR-TE label switched path definition clearly show the label values used (in the Actual Hops section), as follows:

```
[/]
A:admin@PE-1# show router mpls sr-te-lsp "lsp-adj-set" path detail

===============================================================================
MPLS SR-TE LSP lsp-adj-set
Path  (Detail)
===============================================================================
Legend :
    S      - Strict                     L      - Loose
    A-SID  - Adjacency SID              N-SID  - Node SID
    +      - Inherited
===============================================================================
-------------------------------------------------------------------------------
LSP SR-TE lsp-adj-set
Path  path-adj-set
-------------------------------------------------------------------------------
LSP Name    : lsp-adj-set
Path LSP ID      : 37376
From             : 192.0.2.1
To               : 192.0.2.4
Admin State      : Up                       Oper State        : Up
Path Name   : path-adj-set
Path Type        : Primary
Path Admin       : Up                       Path Oper         : Up
Path Up Time     : 0d 00:03:06              Path Down Time    : 0d 00:00:00
Retry Limit      : 0                        Retry Timer       : 30 sec
Retry Attempt    : 0                        Next Retry In     : 0 sec

PathCompMethod   : none                     OperPathCompMethod: none
MetricType       : igp                      Oper MetricType   : igp
LocalSrProt      : preferred                Oper LocalSrProt  : N/A
LabelStackRed    : Disabled                 Oper LabelStackRed: N/A

Bandwidth        : No Reservation           Oper Bandwidth    : 0 Mbps
Hop Limit        : 255                      Oper HopLimit     : 255
Setup Priority   : 7                        Oper SetupPriority: 7
Hold Priority    : 0                        Oper HoldPriority : 0
Inter-area       : N/A

PCE Updt ID      : 0                        PCE Updt State    : None
PCE Upd Fail Code: noError

PCE Report       : Disabled+                Oper PCE Report   : Disabled
```

```
PCE Control      : Disabled                Oper PCE Control  : Disabled

Include Groups   :                         Oper IncludeGroups:
None                                           None
Exclude Groups   :                         Oper ExcludeGroups:
None                                           None
Last Resignal    : n/a

IGP/TE Metric    : 16777215                Oper Metric       : 16777215
Oper MTU         : 1552                    Path Trans        : 1
Degraded         : False
Failure Code     : noError
Failure Node     : n/a
Explicit Hops    :
                    192.0.2.2(L)
                -> 192.0.2.3(S)
                -> 192.0.2.4(L)
Actual Hops      :
    192.0.2.2(192.0.2.2)(N-SID)                  Record Label        : 20002
 -> 192.0.2.3(192.0.2.3)(A-SID)                  Record Label        : 30000
 -> 192.0.2.4(192.0.2.4)(N-SID)                  Record Label        : 20004

BFD Configuration and State
Template         : None                    Ping Interval     : N/A
Enable           : False                   State             : notApplicable
WaitForUpTimer   : 4 sec                   OperWaitForUpTimer: 0 sec
WaitForUpTmLeft  : 0
StartFail Rsn    : N/A

===============================================================================
```

Service configuration

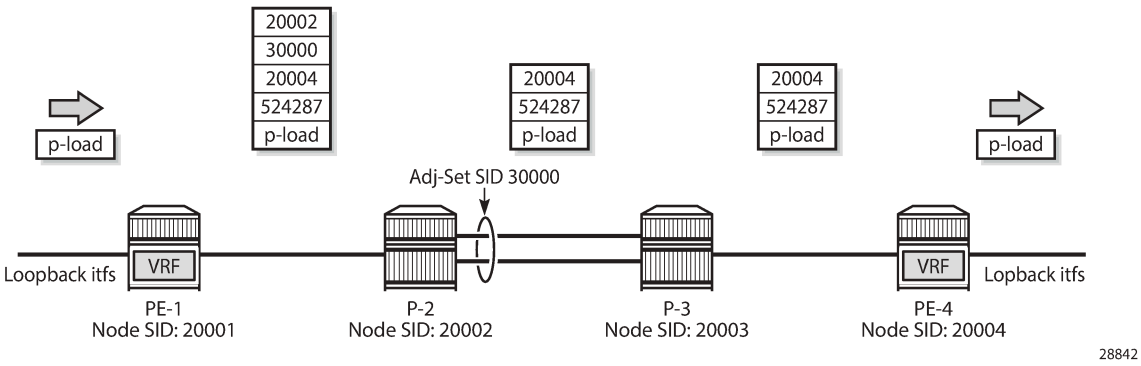
A VPRN service is configured on PE-1 and PE-4, providing multiple loopback interfaces that simulate the TCs. This VPRN is configured to use the SR-TE tunnel defined in the previous section. The configuration on PE-1 is as follows; the configuration on PE-4 is similar.

```
# on PE-1
configure {
    service {
        vprn "svc-1" {
            admin-state enable
            description "runs between PE-1 and PE-4"
            service-id 1
            customer "1"
            autonomous-system 64496
            bgp-ipvpn {
                mpls {
                    admin-state enable
                    route-distinguisher "64496:1"
                    vrf-target {
                        community "target:64496:1"
                    }
                    auto-bind-tunnel {
                        resolution filter
                        resolution-filter {
                            sr-te true
                        }
                    }
                }
            }
```

```
            interface "int_LB_1" {
                loopback true
                ipv4 {
                    primary {
                        address 172.16.14.1
                        prefix-length 32
                    }
                }
            }
            interface "int_LB_2" {
                loopback true
                ipv4 {
                    primary {
                        address 172.16.14.2
                        prefix-length 32
                    }
                }
            }
            interface "int_LB_3" {
                loopback true
                ipv4 {
                    primary {
                        address 172.16.14.3
                        prefix-length 32
                    }
                }
            }
            interface "int_LB_4" {
                loopback true
                ipv4 {
                    primary {
                        address 172.16.14.4
                        prefix-length 32
                    }
                }
            }
            interface "int_LB_5" {
                loopback true
                ipv4 {
                    primary {
                        address 172.16.14.5
                        prefix-length 32
                    }
                }
            }
        }
    }
}
```

Figure 15: MPLS label stack shows the MPLS label stacks on the path from a loopback interface on PE-1 to a loopback interface on PE-4. PE-1 pushes the {20002, 30000, 20004, 524287} label stack to packets received from a loopback interface on PE-1. The bottom of the stack is the {524287} VPN service label. The active label is node SID 20002, so the traffic flow takes the shortest path to P-2, which pops this label. Because {30000} is the label for the adjacency set, this label is popped while spraying the traffic flow across the two links available to P-3. Now the active label is node SID 20004, so the traffic flow takes the shortest path to PE-4; therefore, P-3 swaps {20004} to {20004}. When the traffic flow arrives at PE-4, PE-4 pops the {20004} node SID and the {524287} VPN service label before delivering the traffic to a loopback interface on PE-4.

Figure 15: MPLS label stack



The traffic that is sent in this example is a burst of successive pings (8000) in multiple flows (5) from a loopback interface on PE-1 to the different loopback interfaces on PE-4. So, the traffic flows have a variety of source/destination IP-address pairs. Additionally, for the load to be sprayed across the adjacency set members, load balancing must be enabled. On P-2, this is enabled as follows:

```
# on P-2
configure system load-balancing lsr-load-balancing lbl-ip
```

P-2 hashes the traffic (ping requests) based on the source and destination IP addresses, thereby spraying the traffic across the int-*P-2-P-3-a* interface (on port 1/1/1) and the *int-P-2-P-3-b* interface (on port 1/1/3). P-3 hashes the return traffic (ping responses) similarly across the int-*P-3-P-2-a* interface (on port 1/1/2) and the *int-P-3-P-2-b* interface (on port 1/1/3). Because two links are available, both carry a part of the traffic, as follows. Only the monitoring outcome for P-2 is shown; P-3 has a corresponding monitoring outcome.

```
A:P-2# monitor port 1/1/2 1/1/1 1/1/3 interval 5 repeat 25 absolute

===============================================================================
Monitor statistics for Ports
===============================================================================
                                                    Input                  Output
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
At time t = 0 sec (Base Statistics)
-------------------------------------------------------------------------------
Port 1/1/2
-------------------------------------------------------------------------------
Octets                                             11416                  12136
Packets                                              124                    124
Errors                                                 0                      0

Port 1/1/1
-------------------------------------------------------------------------------
Octets                                             11969                  11703
Packets                                              125                    124
Errors                                                 0                      0

Port 1/1/3
-------------------------------------------------------------------------------
Octets                                             10102                  10102
Packets                                               99                     99
Errors                                                 0                      0
```

```
-------------------------------------------------------------------------------
At time t = 5 sec (Mode: Absolute)
-------------------------------------------------------------------------------
Port 1/1/2
-------------------------------------------------------------------------------
Octets                                          75846                  72206
Packets                                           670                    670
Errors                                              0                      0

Port 1/1/1
-------------------------------------------------------------------------------
Octets                                          44492                  51523
Packets                                           421                    486
Errors                                              0                      0

Port 1/1/3
-------------------------------------------------------------------------------
Octets                                          37859                  30626
Packets                                           352                    286
Errors                                              0                      0

-------------------------------------------------------------------------------

---snip---

-------------------------------------------------------------------------------
At time t = 120 sec (Mode: Absolute)
-------------------------------------------------------------------------------
Port 1/1/2
-------------------------------------------------------------------------------
Octets                                        4734681                4415401
Packets                                         40159                  40159
Errors                                              0                      0

Port 1/1/1
-------------------------------------------------------------------------------
Octets                                        1775114                2654968
Packets                                         16159                  24159
Errors                                              0                      0

Port 1/1/3
-------------------------------------------------------------------------------
Octets                                        2652795                1772795
Packets                                         24127                  16127
Errors                                              0                      0

-------------------------------------------------------------------------------
At time t = 125 sec (Mode: Absolute)
-------------------------------------------------------------------------------
Port 1/1/2
-------------------------------------------------------------------------------
Octets                                        4734964                4415538
Packets                                         40161                  40161
Errors                                              0                      0

Port 1/1/1
-------------------------------------------------------------------------------
Octets                                        1775371                2655251
Packets                                         16162                  24161
Errors                                              0                      0

Port 1/1/3
-------------------------------------------------------------------------------
```

```
Octets                                          2653262                1773262
Packets                                           24131                  16131
Errors                                                0                      0

===============================================================================
A:P-2#
```

The relevant information is available after monitoring all bursts (after 125 seconds): 3 out of 5 flows use the *int-P-2-P-3-a* interface; 2 out of 5 flows use the *int-P-2-P-3-b* interface.

With an additional burst to a loopback interface that is reached above over the *int-P-2-P-3-a* interface, 4 (=3+1) out of now 6 flows use the *int-P-2-P-3-a* interface; the initial 2 out of 6 flows keep on using the *int-P-2-P-3-b* interface, as follows:

```
A:P-2# monitor port 1/1/2 1/1/1 1/1/3 interval 5 repeat 25 absolute

===============================================================================
Monitor statistics for Ports
===============================================================================
                                                    Input                 Output
-------------------------------------------------------------------------------

---snip---

-------------------------------------------------------------------------------
At time t = 125 sec (Mode: Absolute)
-------------------------------------------------------------------------------
Port 1/1/2
-------------------------------------------------------------------------------
Octets                                          5675055                5291655
Packets                                           48120                  48119
Errors                                                0                      0

Port 1/1/1
-------------------------------------------------------------------------------
Octets                                          2651441                3531342
Packets                                           24120                  32120
Errors                                                0                      0

Port 1/1/3
-------------------------------------------------------------------------------
Octets                                          2649909                1769836
Packets                                           24097                  16096
Errors                                                0                      0

===============================================================================
A:P-2#
```

With a further additional burst to a loopback interface that is reached above over the *int-P-2-P-3-b* interface, the initial 4 out of now 7 flows keep on using the *int-P-2-P-3-a* interface; 3 (=2+1) out of 7 flows use the *int-P-2-P-3-b* interface, as follows:

```
A:P-2# monitor port 1/1/2 1/1/1 1/1/3 interval 5 repeat 25 absolute

===============================================================================
Monitor statistics for Ports
===============================================================================
                                                    Input                 Output
-------------------------------------------------------------------------------

---snip---
```

```
-------------------------------------------------------------------------------
At time t = 125 sec (Mode: Absolute)
-------------------------------------------------------------------------------
Port 1/1/2
-------------------------------------------------------------------------------
Octets                                          6618935               6171269
Packets                                           56117                 56115
Errors                                                0                     0

Port 1/1/1
-------------------------------------------------------------------------------
Octets                                          2651175               3531102
Packets                                           24117                 32116
Errors                                                0                     0

Port 1/1/3
-------------------------------------------------------------------------------
Octets                                          3529677               2649844
Packets                                           32094                 24095
Errors                                                0                     0

===============================================================================
A:P-2#
```

Conclusion

By defining adjacency sets in SR-enabled networks, operators can apply load sharing to parallel links between adjacent nodes, thereby optimizing the use of network resources.

Remote Loop-Free Alternate Node Protection

This chapter describes the Remote Loop-Free Alternate Node Protection.

Topics in this chapter include:

Applicability

This chapter was initially written for SR OS Release 16.0.R6, but the MD-CLI in the current edition corresponds to SR OS Release 21.2.R1. Remote Loop-Free Alternate (R-LFA) node protection is supported for IS-IS and OSPF in SR OS Release 16.0.R4 and later. There are no prerequisites for this configuration.

Overview

The Loop-Free Alternates (LFAs) computed following the Remote LFA (R-LFA) specifications in RFC 7490 only guarantee point-to-point link protection by using a repair tunnel. The repair tunnel is a Segment Routed (SR) shortest path between the computing router S and the PQ-node, to ensure that the primary protected link SE is avoided. However, the R-LFA link protection algorithm does not guarantee that the repair path toward the PQ node will avoid the primary next hop router E, and that the traffic emerging from the repair tunnel at the PQ node toward the destination router will avoid the primary next hop router E.

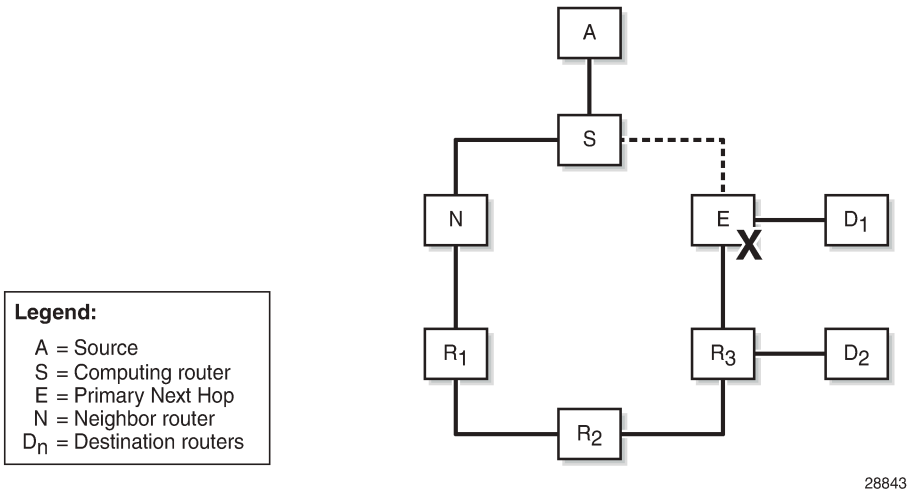
In the remainder of this chapter, SR refers to "Segment Routing", unless specified otherwise. Product and release references, such as 7750 SR and SR OS, continue to refer to "Service Router".

Inequalities for remote LFA node protection

RFC 8102, *Remote LFA node protection*, defines the specifications to protect a path from a source A to a destination D1 or D2, when the primary next hop router E of a computing router S fails; see Figure 16: LFA node protection - topology & denominations. The R-LFA alternate path through a given PQ node to a given destination comprises two path segments:

- path segment from the computing router S to the PQ node (R-LFA alternate next hop)
- path segment from the PQ node to the destination D1 or D2

Figure 16: LFA node protection - topology & denominations



To ensure that an R-LFA alternate path next hop for a given destination provides node protection, none of the path segments may be affected in the event of a failure of the primary next-hop node E. The following four-step algorithm is used to satisfy this requirement:

1. Calculate the node protection extended P-space of router S with respect to the protected node E.
2. Calculate the link protection Q-space of router E with respect to the protected link SE.

   Based on the results of step 1 and 2, a list of one or more candidate PQ-routers is compiled.
3. For each candidate PQ-router, perform an additional forward Shortest Path First (SPF) run to ensure that the path from the PQ-router to the destination router does not traverse the protected router E.
4. If more than one candidate PQ-router satisfies the condition from step 3, router S chooses the PQ-router based on criteria that are specified later in this chapter.

The *node protection extended P-space* is the set of routers Y_i that are reachable from the direct neighbor(s) N of S without traversing protected router E. This excludes the direct neighbors for which there is at least one ECMP path from direct neighbor traversing router E. For a router Y_i to be member of a node protection P-space, the following inequality must be true:

$$cost(N,Y_i) < cost(N,E) + cost(E,Y_i)$$

The *link protection Q-space* is the set of routers that can reach E without traversing the protected link SE, as defined in RFC 7490. This excludes equal cost path routes that traverse the SE link. For a router Y_i to be member of the link protection Q-space, the following inequality must be true:

$$cost(Y_i,E) < cost(Y_i,S) + cost(S,E)$$

If, with respect to router E, a router Y_i is present in the node protection extended P-space and present in the link protection Q-space, it is a candidate PQ node.

Figure 17: Node protecting extended P-space shows the example topology, with metrics and the calculations in table format to determine the node protection extended P-space of router S with respect to the protected node E. Only routers N, R1, and R2 meet the inequality, and therefore belong to the node protecting extended P-space.

Figure 17: Node protecting extended P-space

Router (Yi)	cost(N,Yi)	cost(N,E)	cost(E,Yi)	Inequality met?
N	0	20	20	Yes (0<20+20)
R_1	100	20	120	Yes (100<20+120)
R_2	200	20	200	Yes (200<20+200)
R_3	120	20	100	No (120<20+100)

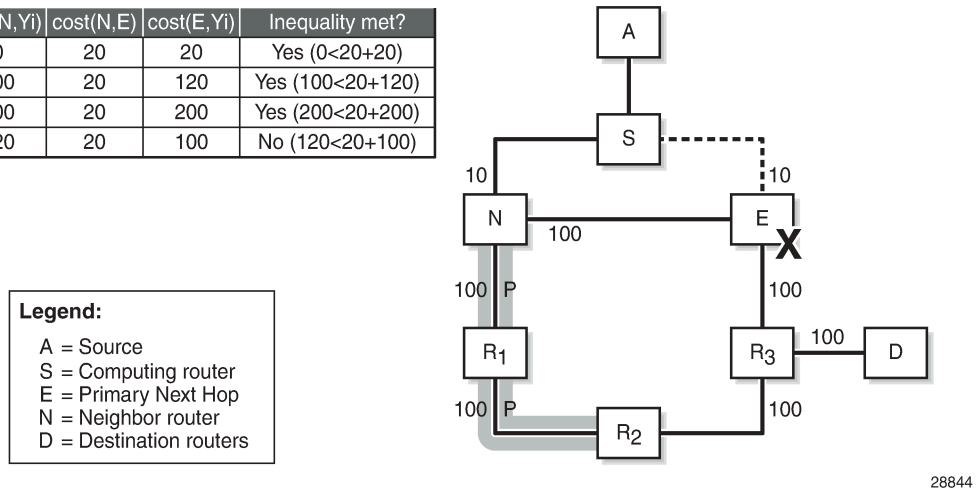
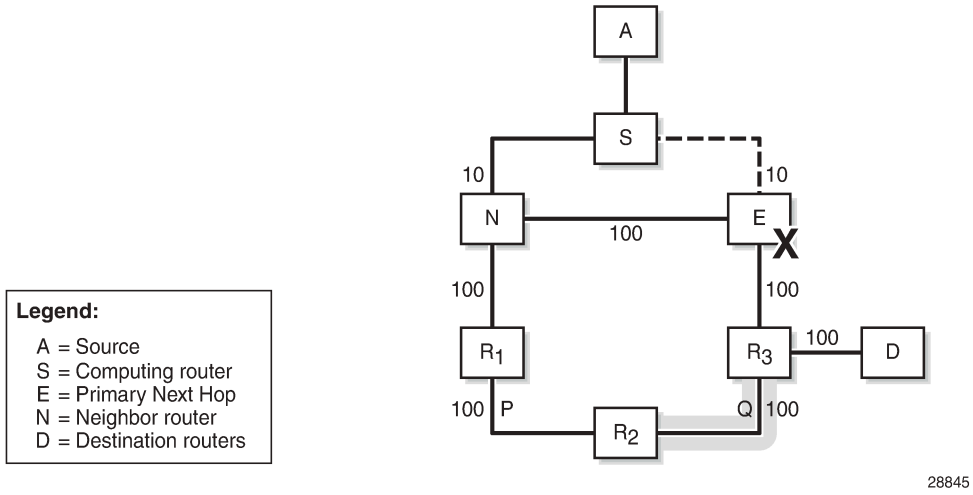


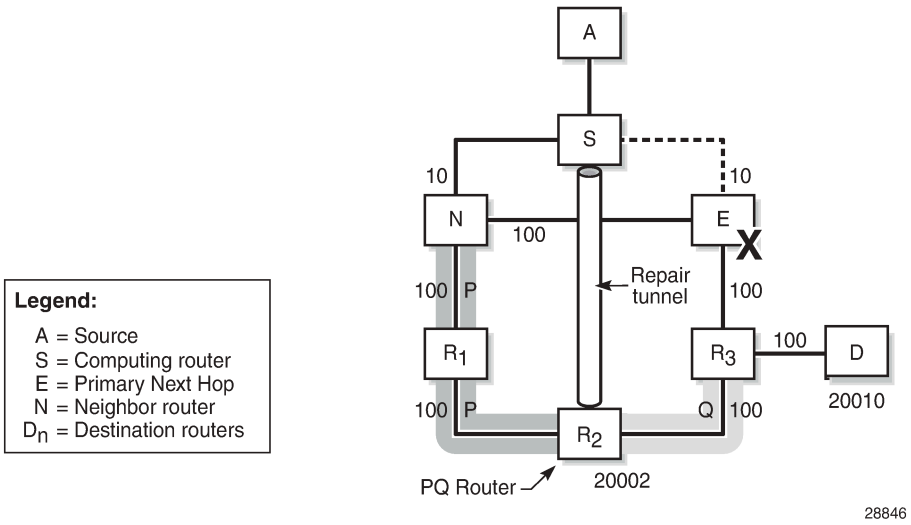
Figure 18: Link protecting Q-space shows the example topology, with metrics and the calculations in table format to determine the link protecting Q-space of router E with respect to the protected link SE. Only routers R2 and R3 meet the inequality, and therefore belong to the link protecting Q-space.

Figure 18: Link protecting Q-space



Candidate PQ routers are routers that belong to the extended P-space and the Q-space. In this example, only R2 is a candidate PQ node; see Figure 19: One candidate PQ-router – repair tunnel.

Figure 19: One candidate PQ-router – repair tunnel



An additional forward SPF run is required to check that the shortest path from the candidate PQ node R2 toward destination D *does not* traverse protected node E. Therefore, the following inequality must be met:

$$cost(PQ_i,D) < cost(PQ_i,E) + cost(E,D)$$

Applied to this topology, the R2-R3-D path does not go via E; therefore, R2 is a valid R-LFA node protection PQ node. The previous inequality evaluates to true, as follows:

$$200 < 200 + 200 \ (True)$$

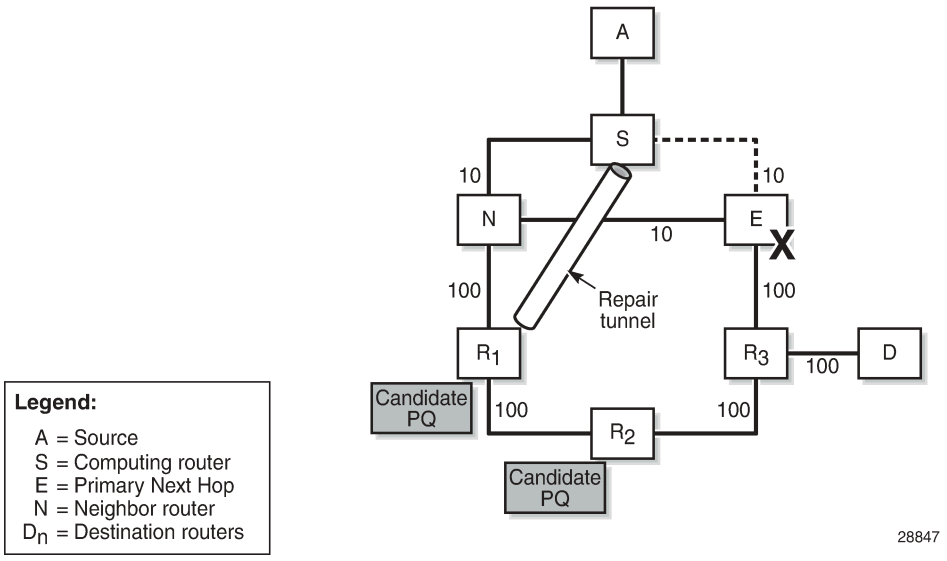
Figure 19: One candidate PQ-router – repair tunnel shows that S constructs a repair tunnel to PQ router R2. To reach destination D from S using the repair tunnel, S pushes a (20002, 20010) label stack, where 20002 and 20010 represent the node SIDs for R2 and D, respectively, while additionally setting the next-hop address to router N. The label 20002 is the first and top label swapped at N and R1, and popped at R2, while 20010 is the second label pushed at R2, swapped at R3, and ultimately popped at D.

In case multiple candidate PQ routers are available, the computing node S selects a PQ router based on the following criteria:

1. lowest IGP path cost from S
2. if multiple PQ routers satisfy (1), S selects the PQ router reachable via the neighbor with the lowest system-ID or router-ID for IS-IS and OSPF, respectively
3. if multiple PQ routers satisfy (1) and (2), S selects the PQ router with the lowest system-ID or router-ID for IS-IS and OSPF, respectively

Figure 20: Two candidate PQ routers – repair tunnel shows an example, with reduced metric between N and E, where R1 and R2 are the candidate PQ routers for protecting router E. In this example, R1 is chosen as the PQ router, because R1 is closer to S than R2. Router S will create an R-LFA repair tunnel for prefixes downstream of R3. To reach those prefixes, the R1 node SID and the D node SID are pushed, with N as the next hop. Prefixes downstream of N, R1, and R2 are unaffected by a failure of E, so they keep using N as their primary next-hop.

Figure 20: Two candidate PQ routers – repair tunnel



LFA and remote LFA interaction

The LFA and remote LFA CLI commands are applied in the OSPF and IS-IS router contexts. The configuration in the IS-IS context is as follows:

```
configure {
    router "Base" {
        isis 0 {
            loopfree-alternate {
                remote-lfa {
                    node-protect
                }
            }
        }
```

Regular LFA is enabled through the **loopfree-alternate** command. Additionally, the **remote-lfa** and **remote-lfa { node-protect }** command can be configured. In other words, by enabling remote LFA, regular LFA is also enabled.

Note:
A remote LFA repair tunnel is only calculated and created if no regular LFA backup next-hop exists. If this is a concern, Topology Independent LFA (TI-LFA) should be enabled; see the Topology-Independent Loop-Free Alternate for Link Protection chapter.

The LFA SPF algorithms are run using the following sequence:

1. A regular LFA is computed for each router and prefix, to provide a backup next-hop per prefix.
2. TI-LFA is computed for all routers and prefixes regardless of the outcome of step 1, and the TI-LFA computed next-hops override the regular LFA next-hops, if TI-LFA is enabled.
3. Remote LFA SPF is only run for the prefixes that are not protected after steps 1 and 2.

As a result, remote LFA next-hops, whether link or node protecting, are only computed and installed when no regular LFA next-hops are available for a given next-hop failure, assuming that TI-LFA is not configured. When the **remote-lfa { node-protect }** command is enabled, the router will prefer a node protect over a link-protect repair tunnel for a given prefix if both are found in the Remote LFA SPF computations.

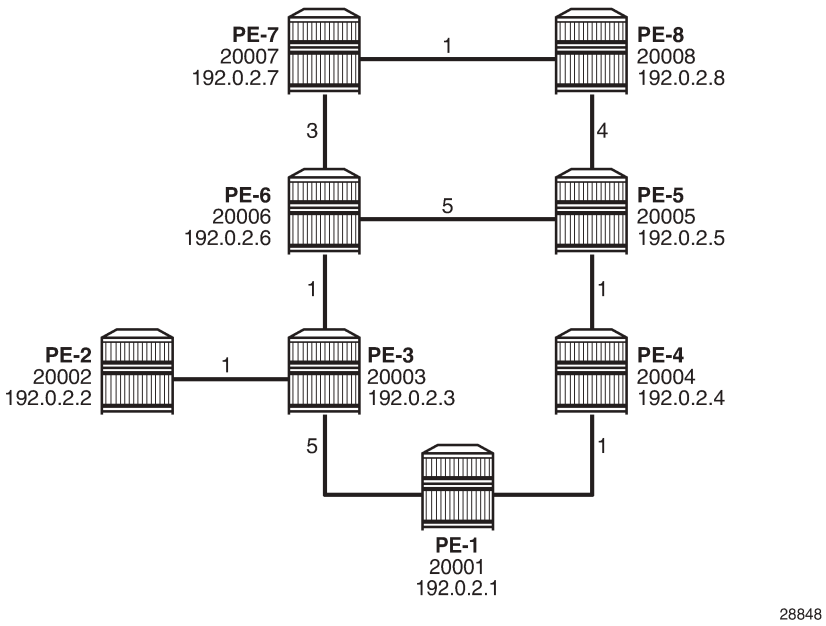
Configuration

Three steps demonstrate the relationship between regular LFA and remote LFA, based on the example topology shown in Figure 21: Example topology. The traffic flow is going from PE-7 to PE-2, and a failure of PE-6 is simulated so that PE-7 is the computing router S, PE-2 is the destination router D, PE-6 is the failing primary next hop E, and PE-8 is the primary backup neighbor N.

The following scenarios are described:

- enable regular LFA on PE-7 — node or link protection cannot be provided for prefixes downstream to PE-6
- enable remote LFA link protection on PE-7 — define the repair tunnel
- enable remote LFA node protection on PE-7 — define the repair tunnel

Figure 21: Example topology



The configuration includes the following:

- Cards, MDAs, ports
- Single stack router interfaces (IPv4 only)
- IS-IS as IGP on the router interfaces. The metrics shown in Figure 21: Example topology are used.
- Segment routing (SR-ISIS) with node SIDs 2000x

The system addresses and the node SIDs for all routers are also shown in Figure 21: Example topology.

Regular LFA

The Segment Routing Global Block (SRGB) is defined consistently across all nodes in the network, as follows:

```
# on all nodes:
configure {
    router "Base" {
        mpls-labels {
            sr-labels {
                start 20000
                end 20099
            }
        }
```

The IS-IS configuration on PE-7 is as follows, and has regular LFA enabled:

```
# on PE-7:
configure {
    router "Base" {
        mpls-labels {
            sr-labels {
                start 20000
                end 20099
            }
        }
        isis 0 {
            admin-state enable
            level-capability 2
            area-address 49.0001.0000
            traffic-engineering true
            advertise-router-capability area
            segment-routing {
                admin-state enable
                prefix-sid-range {
                    global
                }
            }
            loopfree-alternate {
            }
            interface "system" {
                admin-state enable
                ipv4-node-sid {
                    index 7
                }
            }
            interface "int-PE-7-PE-6" {
                admin-state enable
                interface-type point-to-point
                level 2 {
                    metric 3
                }
            }
            interface "int-PE-7-PE-8" {
                admin-state enable
                interface-type point-to-point
                level 2 {
                    metric 1
                }
            }
```

```
        }
```

PE-7 calculates the *regular LFA node protection* for prefixes downstream of PE-6. The shortest path from the primary backup neighbor PE-8 to router PE-2 must be less than the shortest path from the backup neighbor PE-8 node via PE-6, so the inequality becomes:

cost(PE-8,PE-2)) < cost(PE-8,PE-6) + cost(PE-6,PE-2)

(1 + 3 + 1 + 1) < (1 + 3) + (1 + 1) (*False*)

PE-7 calculates the *regular LFA link protection* for the PE-6-PE-7 link for prefixes downstream of PE-6. The shortest path from the primary backup neighbor PE-8 to router PE-2, must be less than the shortest path from the backup neighbor PE-8 via PE-7, so the inequality becomes:

cost(PE-8,PE-2) < cost(PE-8,PE-7) + cost(PE-7,PE-2)

(1 + 3 + 1 + 1) < 1 + (3 + 1 + 1) (*False*)

Because both inequalities are false, PE-7 cannot provide regular LFA PE-6 node protection or regular LFA PE-6-PE-7 link protection.

Remote LFA with link protection

On PE-7, LFA is reconfigured so that *remote LFA with link protection* applies, as follows:

```
# on PE-7:
configure {
    router "Base" {
        isis 0 {
            loopfree-alternate {
                remote-lfa {
                }
            }
        }
```

A repair tunnel will be established, avoiding and protecting the PE-6-PE-7 link, where the endpoint of the repair tunnel is situated on a PQ router.

Figure 22: Link protection extended P-space calculation provides the calculations in table format, along with a graphical representation, to determine the link protecting extended P-space of router PE-7 with respect to the protected PE-6-PE-7 link. Routers PE-1, PE-4, and PE-5 meet the inequality, and therefore belong to the link protecting extended P-space, meaning that they can be reached from backup neighbor PE-8 using an SPF path excluding the PE-6-PE-7 link.

Figure 22: Link protection extended P-space calculation

Yi	cost(PE-8,Yi)	cost(PE-8,PE-7)	cost(PE-7,Yi)	Inequality met?
PE-1	6	1	7	Yes (6 < 1 + 7)
PE-2	6	1	5	No (6 < 1 + 5)
PE-3	5	1	4	No (5 < 1 + 4)
PE-4	5	1	6	Yes (5 < 1 + 6)
PE-5	4	1	5	Yes (4 < 1 + 5)

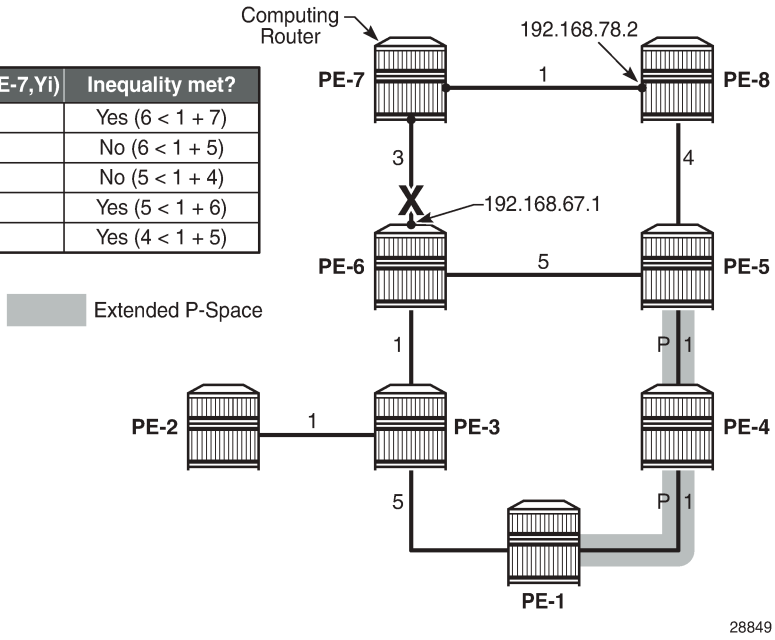


Figure 23: Link protecting Q-space calculation provides the calculations in table format, along with a graphical representation, to determine the link protecting Q-space of router PE-6 with respect to protected PE-7-PE-6 link. Routers PE-1, PE-2, PE-3, PE-4, and PE-5 meet the inequality, and therefore belong to the link protecting Q-space.

Figure 23: Link protecting Q-space calculation

Zi	cost(Zi,PE-6)	cost(Zi,PE-7)	cost(PE-7,PE-6)	Inequality met?
PE-1	6	7	3	Yes (6 < 7 + 3)
PE-2	2	5	3	Yes (2 < 5 + 3)
PE-3	1	4	3	Yes (1 < 4 + 3)
PE-4	6	6	3	Yes (6 < 6 + 3)
PE-5	5	5	3	Yes (5 < 5 + 3)

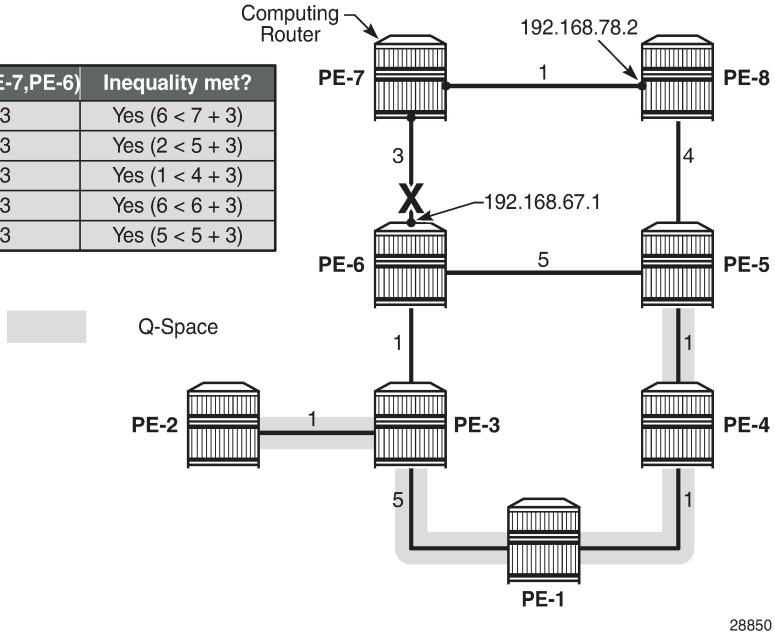
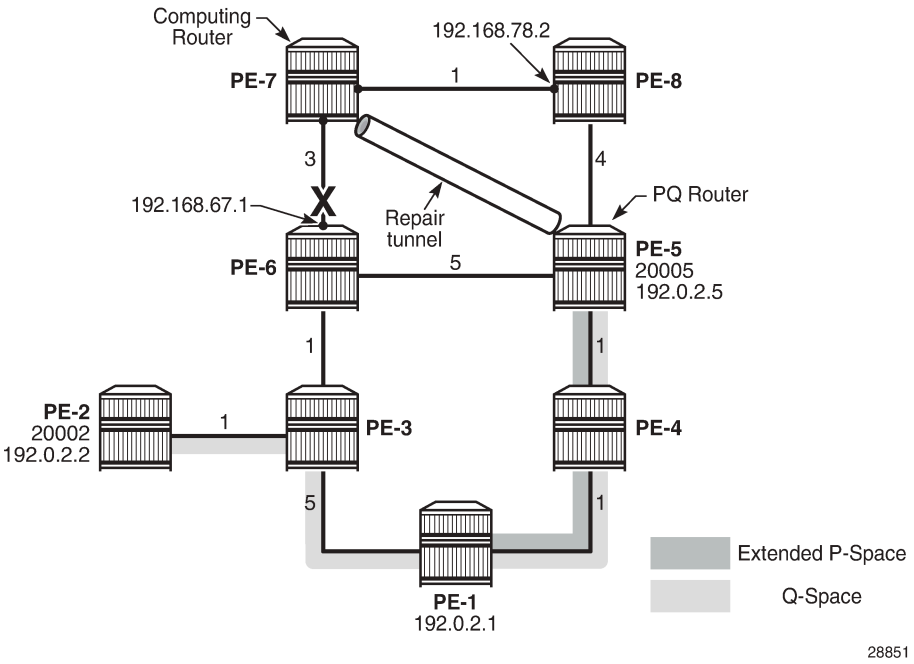


Figure 24: Repair tunnel shows that PE-1, PE-4, and PE-5 are the candidate PQ routers. PE-5 is chosen as the repair tunnel endpoint because of the lowest path cost toward computing node PE-7 (IGP cost from PE-7 to PE-5 = 5). The closest PQ router is chosen to maximize the opportunity for load sharing traffic between the repair tunnel endpoint and the destination router.

Figure 24: Repair tunnel



On the computing node PE-7, the tunnel table for PE-2 destination (192.0.2.2) on IOM 1 shows that 192.168.67.1 is the next hop for the primary path, and that 192.168.78.2 is the next hop for the backup path, as follows. In the normal situation, the PE-7 to PE-2 traffic is routed along the PE-7-PE-6-PE-3-PE-2 path. In case of a PE-7-PE-6 link failure, the traffic on PE-7 node is pushed out with labels 20002 and 20005 to PE-8 (192.168.78.2). The top label is 20005, representing the node SID for PE-5, and 20002 is the label representing the node SID for PE-2.

Note:
Traffic destined for PE-2 and arriving at PE-5 with label 20005 will take the shortest path to PE-2 and therefore will traverse node PE-6.

```
[/]
A:admin@PE-7# show router fp-tunnel-table 1 192.0.2.2/32

===============================================================================
IPv4 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                Protocol           Tunnel-ID
  Lbl
    NextHop                                                   Intf/Tunnel
  Lbl     (backup)
```

```
    NextHop   (backup)
-------------------------------------------------------------------------------
192.0.2.2/32                                SR-ISIS-0           524292
  20002
    192.168.67.1                                                1/1/2
  20002/20005
    192.168.78.2(B)                                             1/1/1
-------------------------------------------------------------------------------
Total Entries : 1
-------------------------------------------------------------------------------
===============================================================================
```

Similar information can be obtained with a **tools dump** command, as follows:

```
[/]
A:admin@PE-7# tools dump router segment-routing tunnel in-label 20002
===============================================================================================================
====
Legend: (B) - Backup Next-hop for Fast Re-Route
        (D) - Duplicate
label stack is ordered from top-most to bottom-most

===============================================================================================================
====
---------------------------------------------------------------------------------------------------------------
---+
 Prefix
 Sid-Type        Fwd-Type       In-Label  Prot-Inst(algoId)
                  Next Hop(s)                                        Out-Label(s) Interface
                                                                           /Tunnel-ID |
---------------------------------------------------------------------------------------------------------------
---+
 192.0.2.2
 Node            Orig/Transit   20002     ISIS-0
                 192.168.67.1                                      20002       int-PE-7-PE-6
               (B)192.168.78.2                                     20005       int-PE-7-PE-8
                                                                   20002
---------------------------------------------------------------------------------------------------------------
---+
No. of Entries: 1
---------------------------------------------------------------------------------------------------------------
---+
```

Another tools command indicates the used LFA type through flags, as follows. Only RLFA link protection applies, and not node protection.

```
[/]
A:admin@PE-7# tools dump router isis sr-database prefix 192.0.2.2 sid 2
===============================================================================
Rtr Base ISIS Instance 0 SR Database
===============================================================================
SID   Label  Prefix          Last-act  Lev MT RtmPref TtmPref Metric  IpNh SrNh
  Mtu    MtuPrim MtuBk  D xL LT Act AdvSystemId    SrErr
-------------------------------------------------------------------------------
2     20002  192.0.2.2       LfaNhops  2   0  18      11      5       1    1
  1556   1564    1564   0 0  R  +R  1920.0000.2002 SR_ERR_OK

-------------------------------------------------------------------------------
No. of Entries: 1
-------------------------------------------------------------------------------
Lev = route level
IpNh = number of IP next-hops
```

```
SrNh = number of SR-tunnel next-hops
D = duplicate pending
xL = exclude from LFA
LT = LFA type (L:LFA, R:RLFA, T:TILFA, n:nodeProtection)
Act = tunnel active state (R:reported, F:failed, +:SR-ack)
===============================================================================
```

Remote LFA with node protection

On PE-7, LFA is reconfigured so that remote LFA with node protection applies, as follows:

```
# on PE-7:
configure {
    router "Base" {
        isis 0 {
            loopfree-alternate {
                remote-lfa {
                    node-protect
                }
            }
        }
```

The general node protecting inequality from the Overview section must be used for defining the node protecting extended P-space. Using the topology from Figure 21: Example topology, the inequality becomes:

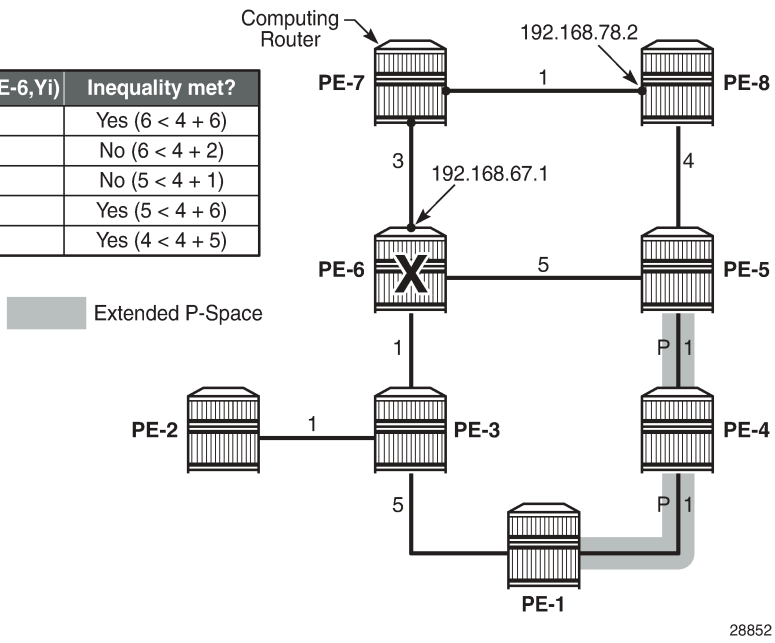
$$cost(N,Y_i) < cost(N,E) + cost(E,Y_i)$$

$$cost(PE\text{-}8,Y_i) < cost(PE\text{-}8,PE\text{-}6) + cost(PE\text{-}6,Y_i)$$

Figure 25: Node protecting extended P-space calculation provides the calculations in table format, along with a graphical representation, to determine the node protecting extended P-space of router PE-7 with respect to protected PE-6 node. Routers PE-1, PE-4, and PE-5 meet the inequality, and therefore belong to the node protecting extended P-space, meaning that they can be reached from backup neighbor PE-8 through an SPF path not passing through node PE-6.

Figure 25: Node protecting extended P-space calculation

Yi	cost(PE-8,Yi)	cost(PE-8,PE-6)	cost(PE-6,Yi)	Inequality met?
PE-1	6	4	6	Yes (6 < 4 + 6)
PE-2	6	4	2	No (6 < 4 + 2)
PE-3	5	4	1	No (5 < 4 + 1)
PE-4	5	4	6	Yes (5 < 4 + 6)
PE-5	4	4	5	Yes (4 < 4 + 5)



The general link protecting inequality from the overview section must be used for defining the Q-space. Using the topology from Figure 21: Example topology, the inequality becomes:

$$cost(Z_i,E) < cost(Z_i,S) + cost(S,E)$$

$$cost(Z_i,PE\text{-}6) < cost(Z_i,PE\text{-}7) + cost(PE\text{-}7,PE\text{-}6)$$

Figure 26: Link protecting Q-space calculation provides the calculations in table format, along with a graphical representation, to determine the link protecting Q-space of router PE-6 with respect to the protected PE-7-PE-6 link. Routers PE-1, PE-2, PE-3, PE-4, and PE-5 meet the inequality, and therefore belong to the link protecting Q-space.

Figure 26: Link protecting Q-space calculation

Zi	cost(Zi,PE-6)	cost(Zi,PE-7)	cost(PE-7,PE-6)	Inequality met?
PE-1	6	7	3	Yes (6 < 7 + 3)
PE-2	2	5	3	Yes (2 < 5 + 3)
PE-3	1	4	3	Yes (1 < 4 + 3)
PE-4	6	6	3	Yes (6 < 6 + 3)
PE-5	5	5	3	Yes (5 < 5 + 3)

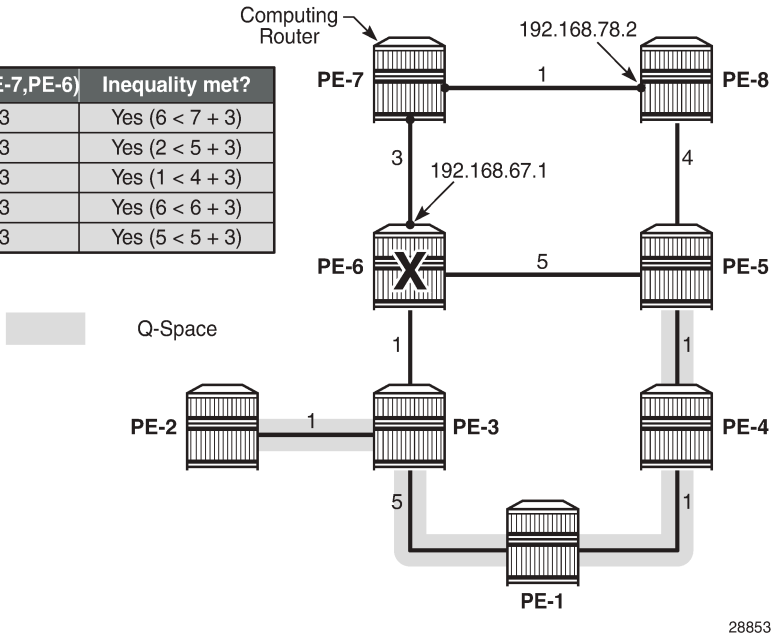
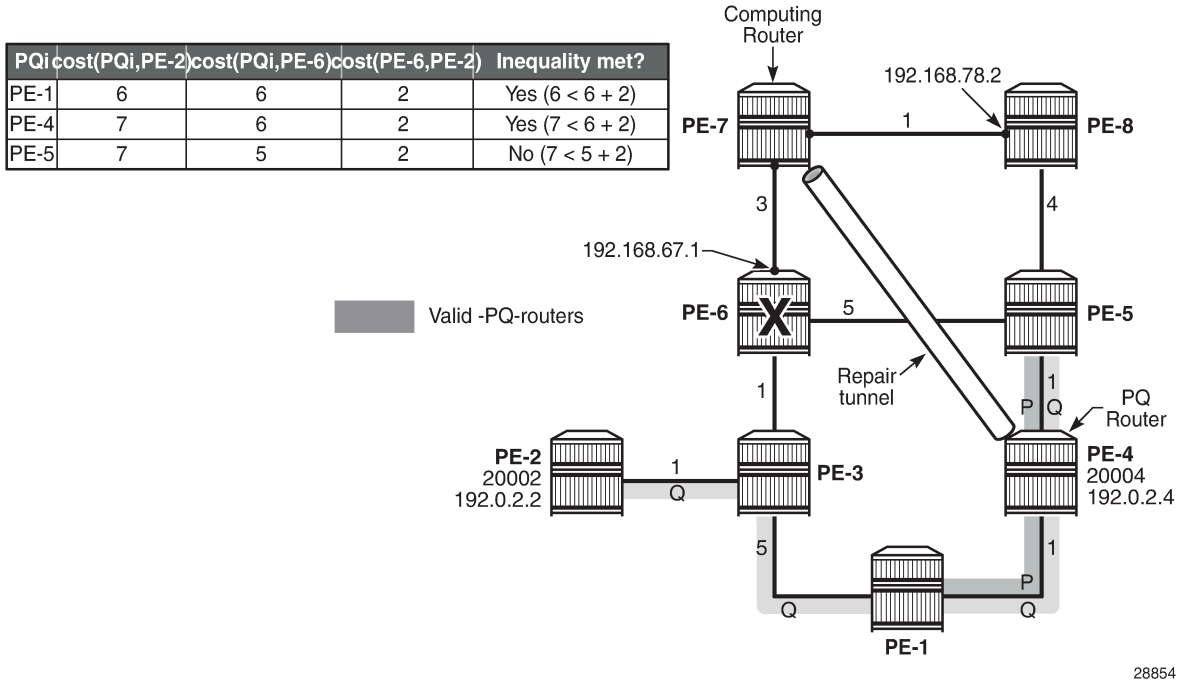


Figure 27: Validating candidate PQ routers - repair tunnel calculation shows that PE-1, PE-4, and PE-5 are the candidate PQ routers for protecting router PE-6. An additional forward SPF run is required for every candidate PQ router, to ensure that the shortest path from that candidate PQ router to destination PE-2 does not traverse the protected router PE-6. The general formula from the Overview section becomes:

$$cost(PQ_i,D) < cost(PQ_i,E) + cost(E,D)$$

$$cost(PQ_i,PE\text{-}2) < cost(PQ_i,PE\text{-}6) + cost(PE\text{-}6,PE\text{-}2)$$

Figure 27: Validating candidate PQ routers - repair tunnel calculation



After validating all three candidate PQ routers, only routers PE-1 and PE-4 are valid for terminating a repair tunnel. The tie-breaker for defining the repair tunnel termination is the lowest IGP path cost from the computing node PE-7 point of view. The cost from PE-7 to PE-4 is lower than the cost from PE-7 to PE-1 (6 < 7), so PE-4 becomes the PQ router.

The tunnel table for destination 192.0.2.2 on IOM 1 shows that 192.168.67.1 is the next hop for the primary path, and that 192.168.78.2 is the next hop for the backup path, as follows. In the normal situation, the PE-7 to PE-2 traffic is routed along the PE-7-PE-6-PE-3-PE-2 path. In case of a PE-6 node failure, the traffic from PE-7 is pushed out to PE-8 (192.168.78.2), with two labels. The label 20004 represents the node SID for PQ node PE-4 and is used as the top (first) label, while 20002 represents the node SID for PE-2 and is used as the second label.

```
[/]
A:admin@PE-7# show router fp-tunnel-table 1 192.0.2.2/32

===============================================================================
IPv4 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                Protocol            Tunnel-ID
  Lbl
    NextHop                                                    Intf/Tunnel
  Lbl     (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
192.0.2.2/32                               SR-ISIS-0            524292
  20002
    192.168.67.1                                               1/1/2
```

```
  20002/20004
    192.168.78.2(B)                                                  1/1/1
-------------------------------------------------------------------------------
Total Entries : 1
-------------------------------------------------------------------------------
===============================================================================
```

Similar information can be obtained with a **tools dump** command, as follows:

```
[/]
A:admin@PE-7# tools dump router segment-routing tunnel in-label 20002
===============================================================================================================
====
Legend: (B) - Backup Next-hop for Fast Re-Route

        (D) - Duplicate

label stack is ordered from top-most to bottom-most

===============================================================================================================
====
---------------------------------------------------------------------------------------------------------------
---+
 Prefix
   |
 Sid-Type        Fwd-Type       In-Label  Prot-Inst(algoId)
  |
              Next Hop(s)                                      Out-Label(s) Interface
                                                                        Tunnel-ID |
---------------------------------------------------------------------------------------------------------------
---+
 192.0.2.2
 Node            Orig/Transit   20002     ISIS-0
                 192.168.67.1                                    20002       int-PE-7-PE-6
               (B)192.168.78.2                                   20004       int-PE-7-PE-8
                                                                 20002
---------------------------------------------------------------------------------------------------------------
---+
No. of Entries: 1
---------------------------------------------------------------------------------------------------------------
---+
```

Another tools command indicates the used LFA type through flags, as follows. RLFA and node protection applies.

```
[/]
A:admin@PE-7# tools dump router isis sr-database prefix 192.0.2.2 sid 2
===============================================================================
Rtr Base ISIS Instance 0 SR Database
===============================================================================
SID   Label  Prefix          Last-act  Lev MT RtmPref TtmPref Metric  IpNh SrNh
  Mtu    MtuPrim MtuBk  D xL LT Act AdvSystemId    SrErr
-------------------------------------------------------------------------------
2     20002  192.0.2.2       LfaNhops  2   0  18      11      5       1    1
  1556   1564    1564   0 0  Rn +R  1920.0000.2002 SR_ERR_OK

-------------------------------------------------------------------------------
No. of Entries: 1
-------------------------------------------------------------------------------
Lev = route level
IpNh = number of IP next-hops
SrNh = number of SR-tunnel next-hops
```

```
D = duplicate pending
xL = exclude from LFA
LT = LFA type (L:LFA, R:RLFA, T:TILFA, n:nodeProtection)
Act = tunnel active state (R:reported, F:failed, +:SR-ack)
===============================================================================
```

The LFA coverage is as follows:

```
[/]
A:admin@PE-7# show router isis sr-lfa-coverage

===============================================================================
Rtr Base ISIS Instance 0 SR LFA Coverage
===============================================================================
MT-ID   SidType      Level Proto LFA       RLFA      TILFA     Coverage
-------------------------------------------------------------------------------
0       node-sid     L2    ipv4  3(42%)    4(57%)    0(0%)     7/7(100%)
0       adj-sid      L2    ipv4  0(0%)     2(100%)   0(0%)     2/2(100%)
===============================================================================
```

Conclusion

Remote LFA Node Protection provides operators the means to create resilient networks, with precalculated backup paths and with improved coverage.

Seamless BFD for SR-TE LSPs

This chapter describes seamless BFD for SR-TE LSPs.

Topics in this chapter include:

Applicability

This chapter was initially written based on SR OS Release 19.10.R1, but the configuration in the current edition is based on SR OS Release 23.3.R3. BFD for RSVP-TE LSPs is supported in SR OS Release 13.0, and later. Seamless BFD for SR-TE LSPs is supported in SR OS Release 19.10.R1, and later.

Overview

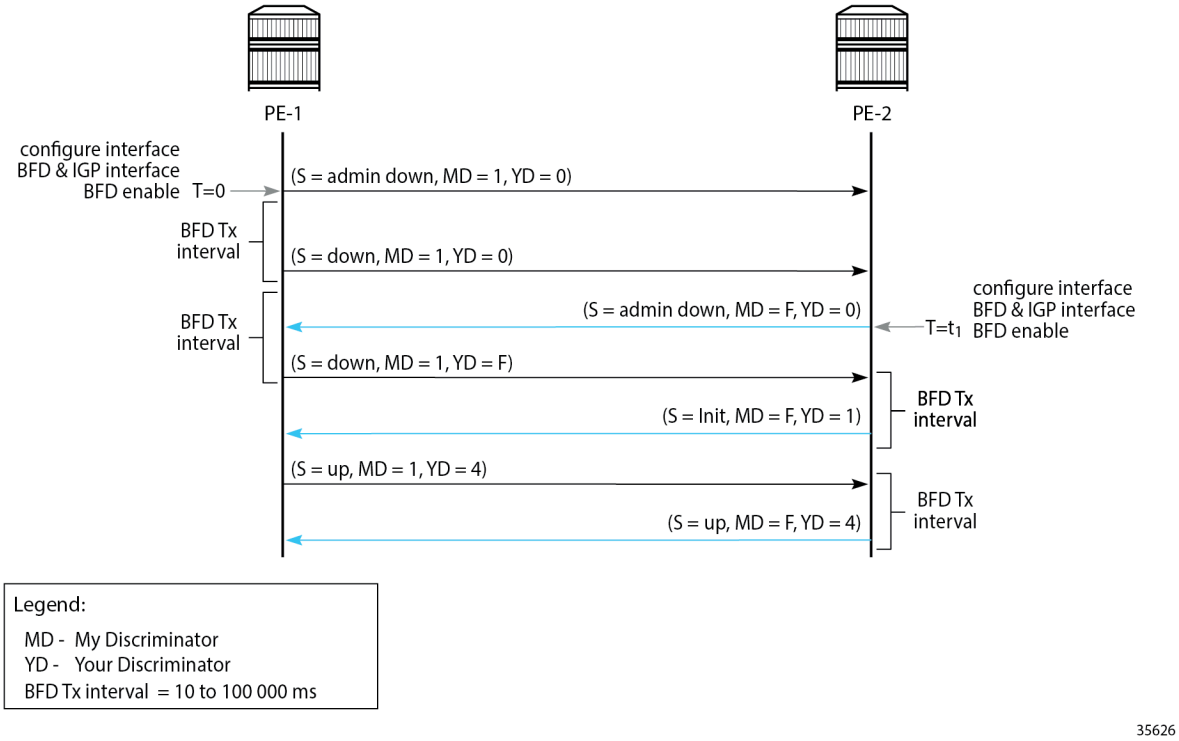
Bidirectional Forwarding Detection (BFD) is widely deployed in IP/MPLS networks to rapidly detect failures in the forwarding path between network elements. In this chapter, a comparison is made between classical BFD and seamless BFD (S-BFD).

Classical BFD

Classical BFD, described in RFC 5880, requires little overhead. However, the handshake mechanism to negotiate and set up two-way BFD sessions between network elements can take several seconds. RFC 5880 specifies two modes of operation: asynchronous mode and on-demand mode. Additionally, the BFD echo function loops back BFD echo packets to the sender.

Classical BFD is applied to the interface. In asynchronous mode, sessions are established. Network elements periodically send BFD control packets to one another. Discriminators are used as a session demultiplexer to distinguish between BFD sessions. The transmitting network element generates a unique non-zero discriminator value, which is exchanged as part of the session handshake establishment. Figure 28: Classical BFD handshake shows the classical BFD handshake for a single hop across an IP link.

Figure 28: Classical BFD handshake



BFD for MPLS LSPs

BFD is supported for RSVP-TE LSPs and for LDP LSPs, as described in the "BFD for RSVP-TE and LDP LSPs" chapter" in the *7450 ESS, 7750 SR, 7950 XRS, and VSR MPLS Advanced Configuration Guide for MD CLI*.

BFD for MPLS LSPs is described in RFC 5884. For continuity checks in MPLS LSPs, BFD packets are transmitted using the MPLS encapsulation, so they share fate with the LSP data path.

BFD is bootstrapped using an LSP ping. An MPLS echo request packet is transmitted along the LSP path, including a BFD discriminator TLV containing the head-end BFD discriminator value. The tail end responds with an echo reply packet, using the IP forwarding path, including the tail-end BFD discriminator value.

Afterward, BFD control packets establish a BFD session between the head end and tail end using the discriminator values from the bootstrap session. The egress LER will send a BFD control packet upon receipt.

Each session has its own pair of discriminators, so multiple discriminators are allocated by the system.

S-BFD for SR-TE LSPs

S-BFD is described in RFC 7880. Unlike classical BFD, S-BFD does not rely on the BFD bootstrapping process (handshake) or session state at the tail end of a session. Instead, when S-BFD is initialized, a pair of discriminators are selected by the system for specific purposes (reflector or initiator). S-BFD minimizes

the time required to establish BFD sessions, which contributes to its seamless operation. S-BFD relies on the fact that the discriminators are already known by the endpoints for each session, either through configuration or advertisement using unicast protocols.

There are two discriminators, one for each end of the BFD/S-BFD session. From the perspective of the S-BFD initiator (or BFD head end) there is a local 'my discriminator' and a remote 'your discriminator'. The 'your discriminator' matches the remote node's local discriminator, which for BFD is allocated to the session endpoint, and for S-BFD is the reflector discriminator.

Terminology

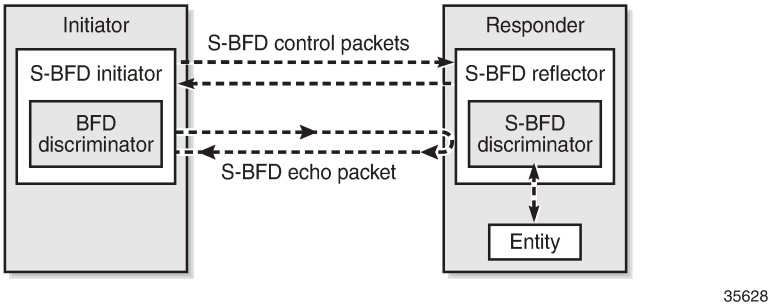
Table 4: RFC 7880 S-BFD terms describes the S-BFD terms, as defined by RFC 7880.

Table 4: RFC 7880 S-BFD terms

S-BFD term	Description
Entity	A function on a network node to which the S-BFD mechanism allows remote network nodes to perform continuity tests. An entity can be abstract (for example, reachability) or specific (for example, IP addresses, router IDs, functions).
S-BFD initiator	An S-BFD session on a network node that performs a continuity test to a remote entity by sending S-BFD packets.
BFD discriminator	An identifier for a BFD session at an endpoint of a BFD session.
Initiator	A network node hosting an S-BFD initiator.
S-BFD reflector	An S-BFD session on a network node that listens for incoming S-BFD control packets to local entities and generates response S-BFD control packets.
Responder	A network node hosting an S-BFD reflector.
S-BFD discriminator	A BFD discriminator allocated for an endpoint of an S-BFD session.

Figure 29: Relationship between S-BFD terms shows the relationship between the S-BFD terms described in Table 4: RFC 7880 S-BFD terms.

Figure 29: Relationship between S-BFD terms



S-BFD implementation in SR OS

Before an application can request the establishment of an S-BFD session, a mapping table of remote discriminators to peer far-end IP addresses must exist. These correspond to the discriminators of the reflector nodes. The mapping can be accomplished in two ways:

- automatically learned (using opaque OSPF or IS-IS routing extensions) or
- statically configured

A single S-BFD discriminator is allocated to a reflector in a router instance. The local reflector S-BFD discriminator is statically configured in the CLI and must be in the range from 524288 to 526335. The S-BFD discriminator must not be the same as any discriminator used for classical BFD.

As per RFC 5884, the destination IP address of explicitly label-switched S-BFD control packets must be chosen from the 127/8 range for IPv4 and the TTL of the IP header must be set to 1. The source IP address is a routable address of the sender.

The initiator node uses the following UDP ports for S-BFD control packets:

- UDP destination port 7784
- UDP source port, which can be any valid port except 7784, as follows:
  - the same UDP source port for all S-BFD control packets to the same reflector
  - different UDP source ports for S-BFD control packets to different reflectors
  - packets with UDP source port 7784 will be discarded by the reflector

The responder node swaps the UDP source and destination port when sending S-BFD control packets back to the initiator node:

- received UDP source port = transmitted UDP destination port
- received UDP destination port = transmitted UDP source port

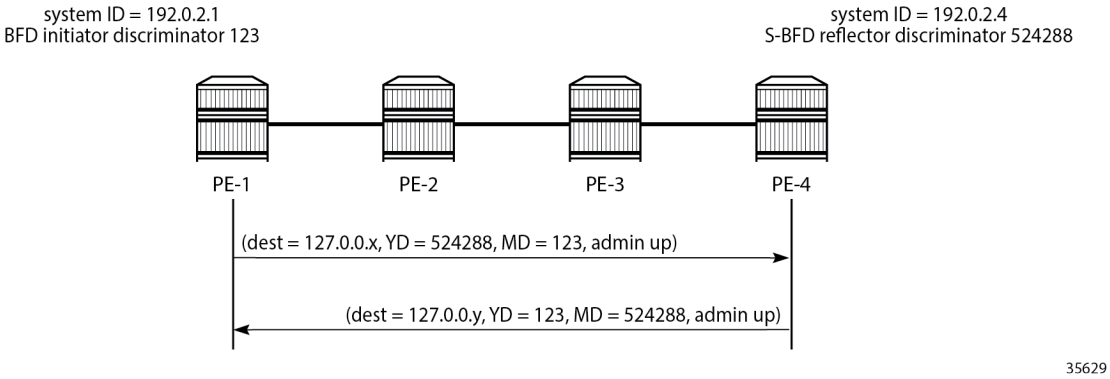
It also exchanges the 'my discriminator' and 'your discrminitator' values in the reflected S-BFD packet.

S-BFD can be applied to SR-TE LSPs and the SR-TE LSP state can depend on the S-BFD session state.

S-BFD session establishment - continuity check

Figure 30: S-BFD session establishment - continuity check shows the continuity check S-BFD control packets between PE-1 and PE-4. On PE-1, the BFD (initiator) discriminator equals 123; on PE-4, the S-BFD (reflector) discriminator equals 524288. Head-end router PE-1 has a mapping table of remote discriminators to far-end IP addresses; for PE-4, the system ID is 192.0.2.4 and the S-BFD discriminator 524288. There is no INIT state in S-BFD. The mapping between the remote discriminators and the far-end IP addresses is required when the BFD return path is routed; when the BFD return path is controlled, no remote discriminators are used.

Figure 30: S-BFD session establishment - continuity check



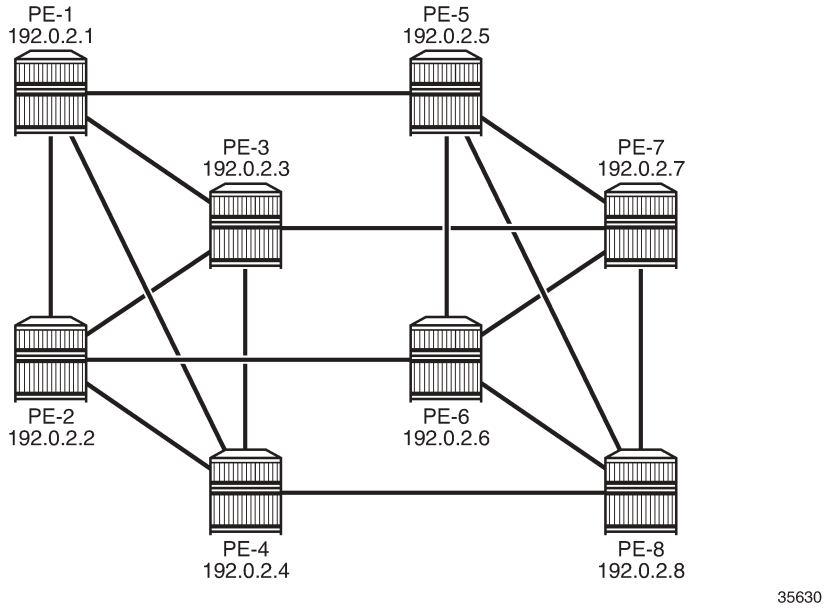
The session initiator node PE-1 generates an S-BFD control packet with destination PE-4 (but with an IP DA from the 127/8 range), YourDiscriminator 524288 (= S-BFD (reflector) discriminator value), MyDiscriminator 123 (= BFD (initiator) discriminator value), and admin state up.

The responder node PE-4 responds to PE-1 with an IP DA from the 127/8 range, YourDiscriminator 123, MyDiscriminator 524288, and admin state up. The admin state of the reflector reflects the configured S-BFD local state.

Configuration

Figure 31: Example topology shows the example topology with eight nodes.

Figure 31: Example topology



The initial configuration includes:

- Cards, MDAs, ports
- Router interfaces
- IS-IS as IGP (alternatively, OSPF can be used) with traffic engineering (TE) enabled
- Segment routing enabled on all nodes
- MPLS and RSVP enabled on all router interfaces

The following will be configured:

- S-BFD for SR-TE LSPs with routed return path between PE-4 and PE-5
- S-BFD for SR-TE LSPs with controlled return path between PE-1 and PE-8

Note:
Even though BFD can use intervals smaller than 1000 ms, the used example setup has its limitations. The nodes in the used example setup are sims and the simulation for CPM-NP or central BFD sessions has the limitation that intervals that are configured with a value smaller than 1000 ms are always negotiated to intervals of 1000 ms. To avoid confusion when the configured intervals differ from the negotiated intervals on sims, a BFD template with intervals of 1000 ms is configured and used in this chapter.

S-BFD for SR-TE LSPs with routed return path

For S-BFD, the S-BFD (reflector) discriminator on the responder (tail-end) node must be known by both end nodes. The mapping between the remote discriminators and the far-end IP addresses can be configured statically or it can be learned dynamically from IGP. On each node, the reflector S-BFD discriminator must be in the range from 524288 to 526335 and the local state must be set to **up**.

Automated S-BFD distribution

In this example, one SR-TE LSP is established between head end PE-4 and tail end PE-5. On tail end PE-5, the global S-BFD configuration is as follows:

```
# on PE-5:
configure {
    bfd {
        seamless-bfd {
            reflector "PE-5" {
                admin-state enable
                discriminator 524292
                local-state up
            }
        }
```

The S-BFD configuration on the other PEs is similar; in this example, it is sufficient to have the global S-BFD configuration on tail end PE-5 only. When the IGP is configured with **advertise-router-capability area** and **traffic-engineering true**, IGP routing protocol extensions provide the encodings to advertise the S-BFD discriminators as opaque information within the IGP link state information. This way, the remote IP addresses and the S-BFD discrimators are automatically mapped.

When PE-4 sets up an SR-TE LSP to PE-5, it will use a BFD discriminator—for example, 3—and S-BFD (reflector) discriminator 524292 for PE-5. For different LSPs toward PE-5, PE-4 will use different BFD discriminators combined with the same S-BFD (reflector) discriminator 524292.

Static S-BFD configuration

If **advertise-router-capability** or **traffic-engineering** are not configured, the S-BFD far-end IP address and its discriminator are statically mapped, as follows. When all SR-TE LSPs have far end PE-5, the mapping for PE-5 is sufficient.

```
# on PE-4:
configure {
    router "Base" {
        bfd {
            seamless-bfd {
                peer 192.0.2.1 {
                    discriminator 524288
                }
                peer 192.0.2.2 {
                    discriminator 524289
                }
                peer 192.0.2.3 {
                    discriminator 524290
                }
                peer 192.0.2.5
                    discriminator 524292
                }
                peer 192.0.2.6 {
                    discriminator 524293
                }
                peer 192.0.2.7 {
                    discriminator 524294
                }
                peer 192.0.2.8 {
                    discriminator 524295
```

```
            }
        }
```

If the initiator receives a valid response from the reflector with an Up state, the initiator declares the S-BFD session as Up.

Note: Traffic engineering in not supported in VPRN or in OSPF3, so S-BFD discriminators cannot be automatically distributed in such cases.

Examples

S-BFD is only supported in CPM-NP on SR OS nodes, so the BFD type must be set to *cpm-np*. SR-TE LSPs can use CPM-NP BFD templates with a transmit and receive interval of minimum 10 ms. However, due to the simulation limitations on the sims in the example topology, the intervals are configured with a value of 1000 ms, as follows:

```
# on PE-4:
configure {
    bfd {
        bfd-template "bfd-cpm-np-1s" {
            receive-interval 1000
            transmit-interval 1000
            type cpm-np
        }
```

On PE-4, the following paths and SR-TE LSPs are configured:

- "LSP-PE-4-PE-5_empty_localCSPF" with primary path "empty", which does not contain any explicit hops
- "LSP-PE-4-PE-5_viaPE-2_localCSPF" with primary path "via-PE-2", which contains 192.0.2.2 as a loose hop
- "LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd" with primary path "via-PE-2" and secondary path "via-PE-3", which contains 192.0.2.3 as a loose hop

Any path computation method can be used. In the following example, the path computation method is local CSPF, as described in the SR-TE LSP Path Computation Using Local CSPF chapter. BFD can be configured per LSP or per path (primary or secondary) in the LSP.

```
# on PE-4:
configure {
    router "Base" {
        mpls {
            path "empty" {
                admin-state enable
            }
            path "via-PE-2" {
                admin-state enable
                hop 10 {
                    ip-address 192.0.2.2
                    type loose
                }
            }
            path "via-PE-3" {
                admin-state enable
                hop 10 {
                    ip-address 192.0.2.3
                    type loose
```

```
            }
        }
        lsp "LSP-PE-4-PE-5_empty_localCSPF" {
            admin-state enable
            type p2p-sr-te
            to 192.0.2.5
            pce-report true
            path-computation-method local-cspf
            max-sr-labels {
                additional-frr-labels 2
            }
            bfd {
                bfd-liveness true
                bfd-template "bfd-cpm-np-1s"
            }
            primary "empty" {
            }
        }
        lsp "LSP-PE-4-PE-5_viaPE-2_localCSPF" {
            admin-state enable
            type p2p-sr-te
            to 192.0.2.5
            pce-report true
            path-computation-method local-cspf
            max-sr-labels {
                additional-frr-labels 2
            }
            primary "via-PE-2" {
                bfd {
                    bfd-liveness true
                    bfd-template "bfd-cpm-np-1s"
                }
            }
        }
        lsp "LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd" {
            admin-state enable
            type p2p-sr-te
            to 192.0.2.5
            pce-report true
            path-computation-method local-cspf
            max-sr-labels {
                additional-frr-labels 2
            }
            bfd {
                bfd-liveness true
                bfd-template "bfd-cpm-np-1s"
            }
            primary "via-PE-2" {
            }
            secondary "via-PE-3" {
                standby true
            }
        }
```

The head-end or initiator node PE-4 learned the S-BFD reflector discriminator for PE-5 (524292), so the BFD control packets can be sent with both a BFD and S-BFD discriminator value. The BFD control packets follow the data path from head end to tail end. The return path is native IP.

The first S-BFD session on initiator node PE-4 gets BFD discriminator 1, the second BFD discriminator 2, and so on. The S-BFD discriminator for PE-5 remains the same: 524292. For "LSP-PE-4-

PE-5_viaPE-2_localCSPF_2nd", with primary and secondary path, two S-BFD sessions are established: one with BFD discriminator 3 and another with BFD discriminator 4, as follows:

```
[/]
A:admin@PE-4# show router bfd seamless-bfd session lsp-name "LSP-PE-4-PE-5_viaPE-2_localCSPF_
2nd" detail

===============================================================================
BFD Session
===============================================================================
Prefix         : 192.0.2.5/32
Local Address  : 192.0.2.4
LSP Name       : LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd
LSP Index      : 65538                    Path LSP ID      : 33792
Fec Type       : srTe
Oper State     : Up                       Protocols        : mplsLsp
Up Time        : 0d 00:07:07              Up Transitions   : 1
Last Down Time : 0d 00:00:01              Down Transitions : 0
                                          Version Mismatch : 0

Forwarding Information

Local Discr    : 3                        Local State      : Up
Local Diag     : 0 (None)
Local Mode     : Demand
Local Min Tx   : 1000                     Local Mult       : 3
Last Sent (ms) : 0                        Local Min Rx     : 0
Type           : cpm-np
Remote Discr   : 524292                   Remote State     : Up
Remote Diag    : 0 (None)                 Remote Mode      : Async
Remote Min Tx  : 1000                     Remote Mult      : 3
Remote C-flag  : 1
Last Recv (ms) : 0                        Remote Min Rx    : 3
-------------------------------------------------------------------------------
Prefix         : 192.0.2.5/32
Local Address  : 192.0.2.4
LSP Name       : LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd
LSP Index      : 65538                    Path LSP ID      : 33794
Fec Type       : srTe
Oper State     : Up                       Protocols        : mplsLsp
Up Time        : 0d 00:07:07              Up Transitions   : 1
Last Down Time : 0d 00:00:01              Down Transitions : 0
                                          Version Mismatch : 0

Forwarding Information

Local Discr    : 4                        Local State      : Up
Local Diag     : 0 (None)
Local Mode     : Demand
Local Min Tx   : 1000                     Local Mult       : 3
Last Sent (ms) : 0                        Local Min Rx     : 0
Type           : cpm-np
Remote Discr   : 524292                   Remote State     : Up
Remote Diag    : 0 (None)                 Remote Mode      : Async
Remote Min Tx  : 1000                     Remote Mult      : 3
Remote C-flag  : 1
Last Recv (ms) : 0                        Remote Min Rx    : 3
===============================================================================
===============================================================================
```

In the preceding **show** command, "Local Discr: 3" and "Local Discr: 4" refer to the BFD discriminator values on the initiator node PE-4, while "Remote Discr: 524292" refers to the S-BFD reflector discriminator value on the responder node PE-5.

The following command shows that the primary path "via-PE-2" goes from PE-4 via PE-2 and PE-1 to PE-5; the secondary path "via-PE-3" goes from PE-4 via PE-3 and PE-7 to PE-5:

```
[/]
A:admin@PE-4# show router mpls sr-te-lsp "LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd" path detail

===============================================================================
MPLS SR-TE LSP LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd
Path  (Detail)
===============================================================================
Legend :
    S      - Strict                     L      - Loose
    A-SID  - Adjacency SID              N-SID  - Node SID
    +      - Inherited
===============================================================================
-------------------------------------------------------------------------------
LSP SR-TE LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd
Path  via-PE-2
-------------------------------------------------------------------------------
LSP Name    : LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd
Path LSP ID      : 33792
From             : 192.0.2.4
To               : 192.0.2.5
Admin State      : Up                       Oper State        : Up
Path Name   : via-PE-2
Path Type        : Primary
Path Admin       : Up                       Path Oper         : Up
---snip---

Explicit Hops    :
                  192.0.2.2(L)
Actual Hops      :
    192.168.24.1(192.0.2.2)(A-SID)               Record Label        : 524286
 -> 192.168.12.1(192.0.2.1)(A-SID)               Record Label        : 524287
 -> 192.168.15.2(192.0.2.5)(A-SID)               Record Label        : 524284

BFD Configuration and State
Template         : None                     Ping Interval     : N/A
Enable           : False                    State             : up
ReturnPathLabel  : None
WaitForUpTimer   : 4 sec                    OperWaitForUpTimer: 4 sec
WaitForUpTmLeft  : 0
StartFail Rsn    : N/A

-------------------------------------------------------------------------------
LSP SR-TE LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd
Path  via-PE-3
-------------------------------------------------------------------------------
LSP Name    : LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd
Path LSP ID      : 33794
From             : 192.0.2.4
To               : 192.0.2.5
Admin State      : Up                       Oper State        : Up
Path Name   : via-PE-3
Path Type        : Standby
Path Admin       : Up                       Path Oper         : Up
---snip---

Explicit Hops    :
```

```
                  192.0.2.3(L)
Actual Hops      :
    192.168.34.1(192.0.2.3)(A-SID)              Record Label        : 524285
 -> 192.168.37.2(192.0.2.7)(A-SID)              Record Label        : 524284
 -> 192.168.57.1(192.0.2.5)(A-SID)              Record Label        : 524286
Srlg             : Disabled               Srlg Disjoint     : False

BFD Configuration and State
Template         : None                   Ping Interval     : N/A
Enable           : False                  State             : up
ReturnPathLabel  : None
WaitForUpTimer   : 4 sec                  OperWaitForUpTimer: 4 sec
WaitForUpTmLeft  : 0
StartFail Rsn    : N/A

===============================================================================
```

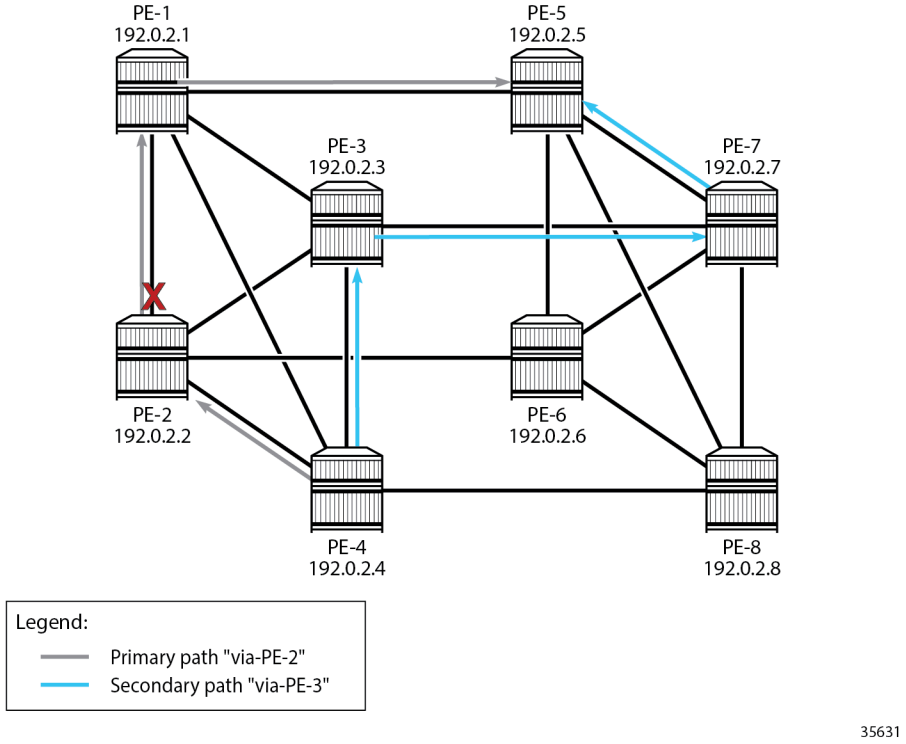
The following OAM LSP trace from PE-4 shows that the path goes via PE-2 and PE-1 to PE-5:

```
[/]
A:admin@PE-4# oam lsp-trace sr-te lsp-name "LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd"
lsp-trace to LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd: 1 hops min, 30 hops max, 176 byte packets
1  192.0.2.2  rtt=2.21ms rc=3(EgressRtr) rsc=3
1  192.0.2.2  rtt=2.85ms rc=8(DSRtrMatchLabel) rsc=2
2  192.0.2.1  rtt=3.03ms rc=3(EgressRtr) rsc=2
2  192.0.2.1  rtt=3.38ms rc=8(DSRtrMatchLabel) rsc=1
3  192.0.2.5  rtt=4.40ms rc=3(EgressRtr) rsc=1
```

S-BFD session down without failure action

Figure 32: Failure on remote link in primary path shows the two paths of "LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd" with a failure on the link between PE-2 and PE-1, which is part of the primary path "via-PE-2". The broken link is remote to the head-end node PE-4. The failure is emulated on PE-2 by disabling the port toward PE-1.

Figure 32: Failure on remote link in primary path



As a result, the BFD session associated with the primary path "via-PE-2" goes down, as follows:

```
[/]
A:admin@PE-4# show router bfd seamless-bfd session lsp-name "LSP-PE-4-PE-5_viaPE-2_localCSPF_
2nd"

===============================================================================
Legend:
  Session Id = Interface Name | LSP Name | Prefix | RSVP Sess Name | Service Id
  wp = Working path   pp = Protecting path
===============================================================================
BFD Session
===============================================================================
Session Id                                        State      Tx Pkts    Rx Pkts
  Rem Addr/Info/SdpId:VcId                      Multipl     Tx Intvl   Rx Intvl
  Protocols                                        Type     LAG Port     LAG ID
  Loc Addr                                                             LAG name
-------------------------------------------------------------------------------
192.0.2.5/32                                      Down          N/A        N/A
  192.0.2.5                                          3         1000          0
  mplsLsp                                       cpm-np          N/A        N/A
  192.0.2.4
192.0.2.5/32                                        Up          N/A        N/A
  192.0.2.5                                          3         1000       1000
  mplsLsp                                       cpm-np          N/A        N/A
  192.0.2.4
-------------------------------------------------------------------------------
No. of BFD sessions: 2
===============================================================================
```

By default, there is no failure action on the BFD session, so the primary path remains up even when the BFD session on that path is down, as follows:

```
[/]
A:admin@PE-4# show router mpls sr-te-lsp "LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd" path detail

===============================================================================
MPLS SR-TE LSP LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd
Path  (Detail)
===============================================================================
Legend :
    S      - Strict                     L      - Loose
    A-SID  - Adjacency SID              N-SID  - Node SID
    +      - Inherited
===============================================================================
-------------------------------------------------------------------------------
LSP SR-TE LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd
Path  via-PE-2
-------------------------------------------------------------------------------
LSP Name    : LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd
Path LSP ID      : 33792
From             : 192.0.2.4
To               : 192.0.2.5
Admin State      : Up                       Oper State        : Up
Path Name   : via-PE-2
Path Type        : Primary
Path Admin       : Up                       Path Oper         : Up
---snip---

Explicit Hops    :
                  192.0.2.2(L)
Actual Hops      :
    192.168.24.1(192.0.2.2)(A-SID)               Record Label        : 524286
 -> 192.168.12.1(192.0.2.1)(A-SID)               Record Label        : 524287
 -> 192.168.15.2(192.0.2.5)(A-SID)               Record Label        : 524284

BFD Configuration and State
Template         : None                     Ping Interval     : N/A
Enable           : False                    State             : down
ReturnPathLabel  : None
WaitForUpTimer   : 4 sec                    OperWaitForUpTimer: 4 sec
WaitForUpTmLeft  : 0
StartFail Rsn    : N/A

-------------------------------------------------------------------------------
---snip---
```

The LSP and its paths remain up and the corresponding SR-TE tunnel in the tunnel table remains unchanged, so the traffic using the LSP will be blackholed. The following tunnel table lists three SR-TE tunnels, corresponding to:

- "LSP-PE-4-PE-5_empty_localCSPF", with next-hop 192.168.48.2 (PE-8)
- "LSP-PE-4-PE-5_viaPE-2_localCSPF", using path "via-PE-2", with next-hop 192.168.24.1
- "LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd", using path "via-PE-2" (the primary path is used, not the secondary), with next-hop 192.168.24.1.

```
[/]
A:admin@PE-4# show router tunnel-table protocol sr-te

===============================================================================
IPv4 Tunnel Table (Router: Base)
```

```
===============================================================================
Destination           Owner     Encap TunnelId  Pref   Nexthop        Metric
   Color
-------------------------------------------------------------------------------
192.0.2.5/32          sr-te     MPLS  655362    8      192.168.48.2   20
192.0.2.5/32          sr-te     MPLS  655363    8      192.168.24.1   30
192.0.2.5/32          sr-te     MPLS  655364    8      192.168.24.1   30
-------------------------------------------------------------------------------
Flags: B = BGP or MPLS backup hop available
       L = Loop-Free Alternate (LFA) hop available
       E = Inactive best-external BGP route
       k = RIB-API or Forwarding Policy backup hop
===============================================================================
```

The OAM LSP ping command using the SR-TE LSP "LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd" fails, as follows:

```
[/]
A:admin@PE-4# oam lsp-ping sr-te lsp-name "LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd"
LSP-PING LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd: 96 bytes MPLS payload
Request timed out.

---- LSP LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd PING Statistics ----
1 packet sent, 0 packets received, 100% packet loss
```

The OAM LSP trace command shows that the LSP trace stops at PE-2 (192.0.2.2):

```
[/]
A:admin@PE-4# oam lsp-trace sr-te lsp-name "LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd"
lsp-trace to LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd: 1 hops min, 30 hops max, 176 byte packets
1  192.0.2.2  rtt=2.15ms rc=3(EgressRtr) rsc=3
1  192.0.2.2  rtt=2.40ms rc=11(DSNoLabelEntry) rsc=2
```

S-BFD session down with failure action

To force a failover to the secondary path or to bring the LSP down when the BFD session goes down, a failure action needs to be configured in the BFD context of the LSP, as follows:

```
# on PE-4:
configure {
    router "Base" {
        mpls {
            lsp "LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd" {
                bfd {
                    bfd-liveness true
                    bfd-template "bfd-cpm-np-1s"
                    failure-action failover-or-down
                    wait-for-up-timer 4    # default; applicable for failure action
                }
```

The failure action **failover-or-down** is the only failure action that is allowed for SR-TE LSPs. An error is raised when attempting to configure failure action **down** or failure action **failover**, as follows:

```
*[ex:/configure router "Base" mpls lsp "LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd" bfd]
A:admin@PE-4# commit
MINOR: MGMT_CORE #3001: configure router "Base" mpls lsp "LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd"
bfd failure-action - 'failure-action down' is not allowed for this LSP type
```

```
*[ex:/configure router "Base" mpls lsp "LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd" bfd]
A:admin@PE-4# commit
MINOR: MGMT_CORE #3001: configure router "Base" mpls lsp "LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd"
bfd failure-action - 'failure-action failover' is not allowed for this LSP type
```

When the failure action is configured, the primary path "via-PE-2" goes down and a failover takes place to the secondary path "via-PE-3" (if available). When no secondary paths are available, the LSP is operational down.

When a link or node fails on the primary path, the BFD state goes down for the primary path. The head-end node switches to the best preference standby that is up. When the LSP retry timer expires, the MPLS module initiates a local CSPF request to calculate a new SR-TE path. When it is possible to calculate a new path meeting the path constraints for the primary path, the new path is added to the SR-TE tunnel, and S-BFD for the primary path is started. S-BFD comes up and the LSP metric is set.

By default, the revert timer is zero, so no reversion to the primary path takes place. However, if the revert timer is configured to a non-zero value, the revert timer starts when the S-BFD session comes up. When the revert timer expires, the active path is reprogrammed from secondary to primary. If **pce-report-enable** is configured, a PCEP status report is sent for each path, so two reports are sent.

The following command shows that the primary path "via-PE-2" is down and the list of actual hops is empty. Therefore, the S-BFD session state is not applicable. The secondary path remains up and the LSP is up.

```
[/]
A:admin@PE-4# show router mpls sr-te-lsp "LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd" path detail

===============================================================================
MPLS SR-TE LSP LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd
Path  (Detail)
===============================================================================
Legend :
    S      - Strict                      L      - Loose
    A-SID  - Adjacency SID               N-SID  - Node SID
    +      - Inherited
===============================================================================
-------------------------------------------------------------------------------
LSP SR-TE LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd
Path  via-PE-2
-------------------------------------------------------------------------------
LSP Name    : LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd
Path LSP ID      : 33796
From             : 192.0.2.4
To               : 192.0.2.5
Admin State      : Up                      Oper State        : Up
Path Name   : via-PE-2
Path Type        : Primary
Path Admin       : Up                      Path Oper         : Down
---snip---

Failure Code     : bfdDown
Failure Node     : 192.0.2.4
Explicit Hops    :
                   192.0.2.2(L)
Actual Hops      :
    No Hops Specified

BFD Configuration and State
Template         : None                    Ping Interval     : N/A
Enable           : False                   State             : notApplicable
ReturnPathLabel  : None
```

```
WaitForUpTimer   : 4 sec                   OperWaitForUpTimer: 4 sec
WaitForUpTmLeft  : 0
StartFail Rsn    : N/A

-------------------------------------------------------------------------------
LSP SR-TE LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd
Path  via-PE-3
-------------------------------------------------------------------------------
LSP Name    : LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd
Path LSP ID      : 33794
From             : 192.0.2.4
To               : 192.0.2.5
Admin State      : Up                      Oper State        : Up
Path Name   : via-PE-3
Path Type        : Standby
Path Admin       : Up                      Path Oper         : Up
---snip---

Failure Code     : noError
Failure Node     : n/a
Explicit Hops    :
                   192.0.2.3(L)
Actual Hops      :
    192.168.34.1(192.0.2.3)(A-SID)               Record Label        : 524285
 -> 192.168.37.2(192.0.2.7)(A-SID)               Record Label        : 524284
 -> 192.168.57.1(192.0.2.5)(A-SID)               Record Label        : 524286
Srlg             : Disabled                Srlg Disjoint     : False

BFD Configuration and State
Template         : None                    Ping Interval     : N/A
Enable           : False                   State             : up
ReturnPathLabel  : None
WaitForUpTimer   : 4 sec                   OperWaitForUpTimer: 4 sec
WaitForUpTmLeft  : 0
StartFail Rsn    : N/A

===============================================================================
```

The tunnel table shows an entry with tunnel ID 655364, which corresponds to "LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd", with next-hop 192.168.34.1 (PE-3):

```
[/]
A:admin@PE-4# show router tunnel-table protocol sr-te

===============================================================================
IPv4 Tunnel Table (Router: Base)
===============================================================================
Destination           Owner     Encap TunnelId  Pref   Nexthop        Metric
   Color
-------------------------------------------------------------------------------
192.0.2.5/32          sr-te     MPLS  655362    8      192.168.48.2   20
192.0.2.5/32          sr-te     MPLS  655363    8      192.168.24.1   30
192.0.2.5/32          sr-te     MPLS  655364    8      192.168.34.1   30
-------------------------------------------------------------------------------
Flags: B = BGP or MPLS backup hop available
       L = Loop-Free Alternate (LFA) hop available
       E = Inactive best-external BGP route
       k = RIB-API or Forwarding Policy backup hop
===============================================================================
```

The OAM LSP trace using "LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd" shows that the active path goes via PE-3 and PE-7 to PE-5, as follows:

```
[/]
A:admin@PE-4# oam lsp-trace sr-te lsp-name "LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd"
lsp-trace to LSP-PE-4-PE-5_viaPE-2_localCSPF_2nd: 1 hops min, 30 hops max, 176 byte packets
1  192.0.2.3  rtt=1.82ms rc=3(EgressRtr) rsc=3
1  192.0.2.3  rtt=2.49ms rc=8(DSRtrMatchLabel) rsc=2
2  192.0.2.7  rtt=4.91ms rc=3(EgressRtr) rsc=2
2  192.0.2.7  rtt=4.18ms rc=8(DSRtrMatchLabel) rsc=1
3  192.0.2.5  rtt=4.11ms rc=3(EgressRtr) rsc=1
```

S-BFD for SR-TE LSPs with controlled return path

In this mode, a controlled return path for BFD reply packets is configured at the initiating node. The reflector function at the far end of the SR-TE LSP is bypassed, so there is no need to configure reflector discriminators for these sessions.

The initiating node pushes an additional MPLS label on S-BFD packets at the bottom of the stack and the BFD session operates in echo mode. The return path label refers to an MPLS binding SID of an SR policy programmed at the far end of the SR-TE LSP. The SR policy can be used to forward BFD reply packets along an explicit TE path back to the initiator, avoiding the IGP shortest path.

It is possible to configure a specific TE return path for each S-BFD session on an SR-TE LSP at the initiating node. The SR policies can have segments lists with different paths, ensuring the BFD reply packets from different LSP paths do not share the same outcome.

In the following example, initiating node PE-1 has three SR-TE LSPs to far end PE-8:

- SR-TE LSP "LSP-PE-1-PE-8_empty_localCSPF" with an empty primary path and return path label 20041
- SR-TE LSP "LSP-PE-1-PE-8_viaPE-2_localCSPF" with primary path "via-PE-2" and return path label 20621
- SR-TE LSP "LSP-PE-1-PE-8_viaPE-2_localCSPF_2nd" with primary path "via-PE-2" and return path label 20621 and secondary path "via-PE-3" and return path label 20051

The configuration of the paths and the SR-TE LSPs on PE-1 is as follows:

```
# on PE-1:
configure {
    bfd {
        bfd-template "bfd-cpm-np-1s" {
            receive-interval 1000
            transmit-interval 1000
            type cpm-np
        }
    }
    router "Base" {
        mpls {
            admin-state enable
            interface "int-PE-1-PE-2" {
            }
            interface "int-PE-1-PE-3" {
            }
            interface "int-PE-1-PE-4" {
            }
            interface "int-PE-1-PE-5" {
```

```
            }
            path "empty" {
                admin-state enable
            }
            path "via-PE-2" {
                admin-state enable
                hop 10 {
                    ip-address 192.0.2.2
                    type loose
                }
            }
            path "via-PE-3" {
                admin-state enable
                hop 10 {
                    ip-address 192.0.2.3
                    type loose
                }
            }
            lsp "LSP-PE-1-PE-8_empty_localCSPF" {
                admin-state enable
                type p2p-sr-te
                to 192.0.2.8
                pce-report true
                path-computation-method local-cspf
                max-sr-labels {
                    additional-frr-labels 2
                }
                bfd {
                    bfd-liveness true
                    bfd-template "bfd-cpm-np-1s"
                    failure-action failover-or-down
                    return-path-label 20041
                }
                primary "empty" {
                }
            }
            lsp "LSP-PE-1-PE-8_viaPE-2_localCSPF" {
                admin-state enable
                type p2p-sr-te
                to 192.0.2.8
                pce-report true
                path-computation-method local-cspf
                max-sr-labels {
                    additional-frr-labels 2
                }
                bfd {
                    failure-action failover-or-down
                }
                primary "via-PE-2" {
                    bfd {
                        bfd-liveness true
                        bfd-template "bfd-cpm-np-1s"
                        return-path-label 20621
                    }
                }
            }
            lsp "LSP-PE-1-PE-8_viaPE-2_localCSPF_2nd" {
                admin-state enable
                type p2p-sr-te
                to 192.0.2.8
                pce-report true
                path-computation-method local-cspf
                max-sr-labels {
                    additional-frr-labels 2
```

```
                }
                bfd {
                    failure-action failover-or-down
                }
                primary "via-PE-2" {
                    bfd {
                        bfd-liveness true
                        bfd-template "bfd-cpm-np-1s"
                        return-path-label 20621
                    }
                }
                secondary "via-PE-3" {
                    standby true
                    bfd {
                        bfd-liveness true
                        bfd-template "bfd-cpm-np-1s"
                        return-path-label 20051
                    }
                }
            }
```

The return path labels correspond to binding SIDs in SR policies on PE-8, as follows:

```
# on PE-8:
configure {
    router "Base" {
        mpls-labels {
            sr-labels {
                start 32000
                end 32999
            }
            reserved-label-block "SRLB1" {
                start-label 20000
                end-label 21999
            }
        }
        segment-routing {
            sr-policies {
                admin-state enable
                reserved-label-block "SRLB1"
                static-policy "SR-static-policy-PE-4-PE-1" {
                    admin-state enable
                    color 810
                    endpoint 192.0.2.1
                    head-end local
                    binding-sid 20041
                    distinguisher 10020041
                    segment-list 1 {
                        admin-state enable
                        segment 1 {
                            mpls-label 32004    # node SID for PE-4
                        }
                        segment 2 {
                            mpls-label 32001    # node SID for PE-1
                        }
                    }
                }
                static-policy "SR-static-policy-PE-5-PE-1" {
                    admin-state enable
                    color 820
                    endpoint 192.0.2.1
                    head-end local
                    binding-sid 20051
```

```
                    distinguisher 10020051
                    segment-list 1 {
                        admin-state enable
                        segment 1 {
                            mpls-label 32005    # node SID for PE-5
                        }
                        segment 2 {
                            mpls-label 32001    # node SID for PE-1
                        }
                    }
                }
                static-policy "SR-static-policy-PE-6-PE-2-PE-1" {
                    admin-state enable
                    color 830
                    endpoint 192.0.2.1
                    head-end local
                    binding-sid 20621
                    distinguisher 10020621
                    segment-list 1 {
                        admin-state enable
                        segment 1 {
                            mpls-label 32006    # node SID for PE-6
                        }
                        segment 2 {
                            mpls-label 32002    # node SID for PE-2
                        }
                        segment 3 {
                            mpls-label 32001    # node SID for PE-1
                        }
                    }
                }
            }
        }
```

The tunnel table on PE-8 contains three SR-policy tunnels to PE-1:

```
[/]
A:admin@PE-8# show router tunnel-table protocol sr-policy

===============================================================================
IPv4 Tunnel Table (Router: Base)
===============================================================================
Destination           Owner     Encap TunnelId  Pref   Nexthop        Metric
   Color
-------------------------------------------------------------------------------
192.0.2.1/32          sr-policy MPLS  917506    14     192.0.2.4      0
   810
192.0.2.1/32          sr-policy MPLS  917507    14     192.0.2.5      0
   820
192.0.2.1/32          sr-policy MPLS  917508    14     192.0.2.6      0
   830
-------------------------------------------------------------------------------
Flags: B = BGP or MPLS backup hop available
       L = Loop-Free Alternate (LFA) hop available
       E = Inactive best-external BGP route
       k = RIB-API or Forwarding Policy backup hop
===============================================================================
```

Four S-BFD sessions are up between PE-1 and PE-8:

```
[/]
A:admin@PE-1# show router bfd session
```

```
===============================================================================
Legend:
  Session Id = Interface Name | LSP Name | Prefix | RSVP Sess Name | Service Id
  wp = Working path   pp = Protecting path
===============================================================================
BFD Session
===============================================================================
Session Id                                        State      Tx Pkts    Rx Pkts
  Rem Addr/Info/SdpId:VcId                      Multipl     Tx Intvl   Rx Intvl
  Protocols                                        Type     LAG Port     LAG ID
  Loc Addr                                                             LAG name
-------------------------------------------------------------------------------
192.0.2.8/32                                         Up          N/A        N/A
  192.0.2.8                                           3         1000       1000
  mplsLsp                                         cpm-np          N/A        N/A
  192.0.2.1
192.0.2.8/32                                         Up          N/A        N/A
  192.0.2.8                                           3         1000       1000
  mplsLsp                                         cpm-np          N/A        N/A
  192.0.2.1
192.0.2.8/32                                         Up          N/A        N/A
  192.0.2.8                                           3         1000       1000
  mplsLsp                                         cpm-np          N/A        N/A
  192.0.2.1
192.0.2.8/32                                         Up          N/A        N/A
  192.0.2.8                                           3         1000       1000
  mplsLsp                                         cpm-np          N/A        N/A
  192.0.2.1
-------------------------------------------------------------------------------
No. of BFD sessions: 4
===============================================================================
```

When the SR policies on PE-8 are down, the corresponding BFD sessions on PE-1 go down.

On PE-1, SR-TE LSP "LSP-PE-1-PE-8_viaPE-2_localCSPF_2nd" has a primary path and a standby secondary path. The local discriminator for the primary path is 7; for the secondary path 5. No remote discriminators are used when the return path corresponds to an SR policy, so the remote discrimators equal zero. The return path label is the binding SID of the SR policy in the far end node.

```
[/]
A:admin@PE-1# show router bfd seamless-bfd session lsp-name "LSP-PE-1-PE-8_viaPE-2_localCSPF_
2nd" detail

===============================================================================
BFD Session
===============================================================================
Prefix         : 192.0.2.8/32
Local Address  : 192.0.2.1
LSP Name       : LSP-PE-1-PE-8_viaPE-2_localCSPF_2nd
LSP Index      : 65538                     Path LSP ID      : 54784
Fec Type       : srTe
Return Path    : 20621
Oper State     : Up                        Protocols        : mplsLsp
Up Time        : 0d 00:05:18               Up Transitions   : 1
Last Down Time : 0d 00:00:01               Down Transitions : 0
                                           Version Mismatch : 0

Forwarding Information

Local Discr    : 7                         Local State      : Up
Local Diag     : 0 (None)
Local Mode     : Demand
Local Min Tx   : 1000                      Local Mult       : 3
```

```
Last Sent (ms) : 0                        Local Min Rx     : 1000
Type           : cpm-np
Remote Discr   : 0                        Remote State     : Up
Remote Diag    : 0 (None)                 Remote Mode      : Demand
Remote Min Tx  : 1000                     Remote Mult      : 3
Remote C-flag  : 1
Last Recv (ms) : 0                        Remote Min Rx    : 1000
===============================================================================
Prefix         : 192.0.2.8/32
Local Address  : 192.0.2.1
LSP Name       : LSP-PE-1-PE-8_viaPE-2_localCSPF_2nd
LSP Index      : 65538                    Path LSP ID      : 54786
Fec Type       : srTe
Return Path    : 20051
Oper State     : Up                       Protocols        : mplsLsp
Up Time        : 0d 00:05:27              Up Transitions   : 1
Last Down Time : 0d 00:00:01              Down Transitions : 0
                                          Version Mismatch : 0

Forwarding Information

Local Discr    : 5                       Local State      : Up
Local Diag     : 0 (None)
Local Mode     : Demand
Local Min Tx   : 1000                     Local Mult       : 3
Last Sent (ms) : 0                        Local Min Rx     : 1000
Type           : cpm-np
Remote Discr   : 0                        Remote State     : Up
Remote Diag    : 0 (None)                 Remote Mode      : Demand
Remote Min Tx  : 1000                     Remote Mult      : 3
Remote C-flag  : 1
Last Recv (ms) : 0                        Remote Min Rx    : 1000
===============================================================================
===============================================================================
```

Conclusion

Seamless BFD for SR-TE LSPs allows fast connectivity checking of the data plane of the LSP. This can be used to trigger fast failover from the currently active to a standby path.

Segment Routing – Traffic Engineered Tunnels

This chapter provides information about Segment Routing – Traffic Engineered Tunnels.

Topics in this chapter include:

Applicability

This chapter was initially written for SR OS Release 14.0.R7, but the MD-CLI in the current edition corresponds to SR OS Release 25.10.R2.

Overview

Segment Routing (SR) is described in the chapter Segment Routing with IS-IS Control Plane, where the advertisement of node prefix segment identifiers (SIDs) cause the automatic creation of ECMP-aware shortest path MPLS tunnels on each SR-aware router. Each node prefix SID is a globally unique value and becomes an MPLS label in the MPLS data plane. The label is advertised and learned by each SR-capable router using control plane extensions to the IS-IS and OSPF protocols.

It is also possible to create source-routed traffic-engineered end-to-end segment routing paths, where routing constraints such as strict or loose hops can be used to determine a data path to be taken through a network.

These are known as Segment Routing Traffic Engineered (SR-TE) Label Switched Paths (LSPs) and use the same command line construct as that used in configuring RSVP-TE LSPs. However, SR-TE LSPs differ in that there is no mid-point state; each intermediate and tail-end router is unaware of the presence of the LSP because there is no signaling protocol used to create the path. The path can be computed locally by the ingress PE or by offloading the path computation to an external controller.

If a packet is forwarded through the SR tunnel, each router along the path reads the top label and forwards the packet according to the SR tunnel table entry for that label.

This chapter describes the configuration of SR-TE LSPs with locally-computed source-routed paths and how they can be used in the data plane of Layer 2 and Layer 3 services. In the cases described, an SR-TE LSP containing a number of strict or loose hops is created at the head-end router and used to construct an LSP by translating the IP addresses configured in the MPLS path to an SID. This results in an MPLS path with state at the head end only, comprising a stack of SIDs, where each SID is an MPLS label.

In this chapter, OSPF is used to advertise the SIDs and a set of extensions to OSPF have been defined, which require additional configuration on each network router.

The LSP is instantiated—the state is operationally "up"—and a tunnel table entry is created that is owned by the SR-TE protocol. Any data packet that is resolved to use the resulting tunnel has the label stack imposed at the head-end router and is forwarded out of the appropriate next-hop interface. This interface is determined by the topmost label in the stack.

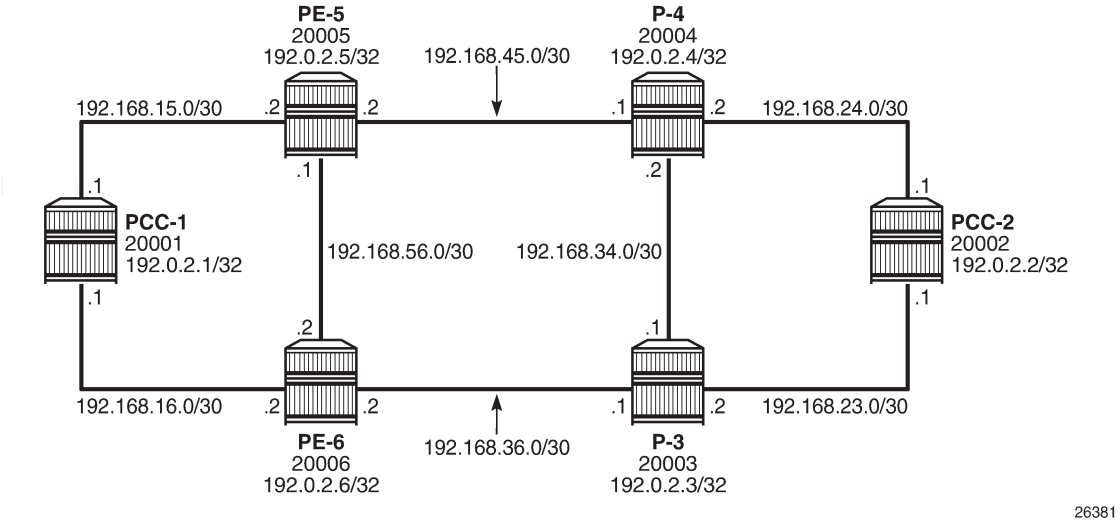
If the label is a node SID, the outgoing interface is determined by the IGP—the shortest path to the router that the node SID represents.

If the label is a local adjacency SID, the outgoing interface is the local interface for which this SID is generated by the IGP.

The segments referenced can be a prefix segment, such as a node segment or an adjacency segment, which represents a specific adjacency between two nodes. The SIDs are used as MPLS labels.

In the following configuration examples, the LSP path is created at the head-end router, and computed by translating a list of hops containing IP addresses into a list of SIDs, by examining the OSPF TE database. The head-end router is referred to as a Path Computation Client (PCC). Figure 33: Segment routing network schematic shows the example topology used, and a pair of bidirectional connected SR-TE LSPs between PCC-1 and PCC-2 is configured to illustrate SR-TE LSPs. All interfaces between PCC-1 and its neighbors have the OSPF metric set to 1000. Similarly, for PCC-2, the OSPF metric is also set to 1000 between itself and its neighbors. The OSPF metric on router interfaces between the core routers P-3, P-4, PE-5, and PE-6 are set to 100.

Figure 33: Segment routing network schematic



Configuration

MPLS label range

The MPLS label range must be configured. This represents the Segment Routing Global Block (SRGB) from which node SIDs are allocated. The choice of SRGB in this example is the same as that chosen in the

chapter Segment Routing with IS-IS Control Plane, where the label block is the same for each router. The SRGB is a contiguous range within the dynamic range 18432 to 524287, as shown in the following output:

```
[/]
A:admin@PCC-1# show router mpls-labels label-range

===============================================================================
Label Ranges
===============================================================================
Label Type      Start Label End Label   Aging       Available   Total
-------------------------------------------------------------------------------
Static          32          18431       -           18400       18400
Dynamic         18432       524287      0           505856      505856
    Seg-Route   0           0           -           0           0
===============================================================================
```

In this example, a range of 1000 labels is chosen. For operational simplicity, Nokia recommends that the same label range is chosen for each router. However, this is not an explicit requirement.

A label range of 20000 to 20999 for SR is configured with the following command:

```
# on all nodes:
configure {
    router "Base" {
        mpls-labels {
            sr-labels {
                start 20000
                end 20999
            }
        }
```

When the SRGB label range has been configured, the MPLS label range looks as follows:

```
[/]
A:admin@PCC-1# show router mpls-labels label-range

===============================================================================
Label Ranges
===============================================================================
Label Type      Start Label End Label   Aging       Available   Total
-------------------------------------------------------------------------------
Static          32          18431       -           18400       18400
Dynamic         18432       524287      0           504856      505856
    Seg-Route   20000       20999       -           0           1000
===============================================================================
```

Global OSPF configuration

The first step is to configure OSPF on each router, as shown in Figure 33: Segment routing network schematic. All router interfaces are members of a single backbone area: area 0.0.0.0.

The configuration for PCC-1 to enable OSPF is:

```
# on PCC-1:
configure {
    router "Base" {
        ospf 0 {
            admin-state enable
            area 0.0.0.0 {
```

```
                interface "int-PCC-1-PE-5" {
                    interface-type point-to-point
                    metric 1000
                }
                interface "int-PCC-1-PE-6" {
                    interface-type point-to-point
                    metric 1000
                }
                interface "system" {
                }
            }
        }
```

The configuration for all other nodes is the same, apart from the IP addresses. The IP addresses can be derived from Figure 33: Segment routing network schematic.

For each router to be segment-routing capable, additional configuration within the OSPF context is required. For PCC-1, this is as follows:

```
# on PCC-1:
configure {
    router "Base" {
        ospf 0 {
            admin-state enable
            advertise-router-capability area
            traffic-engineering true
            segment-routing {
                admin-state enable
                prefix-sid-range {
                    global
                }
            }
            area 0.0.0.0 {
                interface "system" {
                     node-sid {
                        label 20001
                    }
                }
```

The router capability is enabled using the **advertise-router-capability area** command, which defines the flooding scope of the opaque LSA used for this purpose as area. Traffic engineering is also enabled.

Also, MPLS and RSVP must be enabled on each router interface to ensure that OSPF opaque LSAs are generated.

A node SID is manually configured as a label, equivalent to the absolute node SID value. It is possible to configure the node SID as an index. Indexing is described in the Segment Routing with IS-IS Control Plane chapter.

Finally, segment routing is enabled, along with the **prefix-sid-range** command that states that the node prefix SID values of all routers within the network is within the range of the global block.

The value of the **prefix-sid-range** must be the same for all routers; in this case, the range is always 1000.

The following output taken from PCC-1 shows the prefix SIDs configured on the routers in the network and advertised using OSPF. The output is similar for all routers in the network.

```
/]
A:admin@PCC-1# show router ospf prefix-sids

===============================================================================
Rtr Base OSPFv2 Instance 0 Prefix-Sids
```

```
===============================================================================
Prefix                          Area            RtType     SID        Shared
                                Adv-Rtr         SRMS       Flags
-------------------------------------------------------------------------------
192.0.2.1/32                    0.0.0.0         INTRA-AREA 1          No
                                192.0.2.1       N          NnP
192.0.2.2/32                    0.0.0.0         INTRA-AREA 2          N.A.
                                192.0.2.2       N          NnP
192.0.2.3/32                    0.0.0.0         INTRA-AREA 3          N.A.
                                192.0.2.3       N          NnP
192.0.2.4/32                    0.0.0.0         INTRA-AREA 4          N.A.
                                192.0.2.4       N          NnP
192.0.2.5/32                    0.0.0.0         INTRA-AREA 5          N.A.
                                192.0.2.5       N          NnP
192.0.2.6/32                    0.0.0.0         INTRA-AREA 6          N.A.
                                192.0.2.6       N          NnP
-------------------------------------------------------------------------------
No. of Prefix/SIDs: 6
SRMS      : Y/N = prefix SID advertised by SR Mapping Server (Y) or not (N)
            S   = SRMS prefix SID is selected to be programmed
SID Flags : N = Node-SID
            nP = no penultimate hop POP
            M = Mapping server
            E = Explicit-Null
            V = Prefix-SID carries a value
            L = value/index has local significance
            I = Inter Area flag
            A = Attached flag
            B = Backup flag
Shared    : Yes  = local shared Node-SID
            No   = not a local shared Node-SID
            N.A. = not applicable for Remote prefix-sid
===============================================================================
```

The prefix SID for each node is displayed as an index; for example, 1. The absolute value of the node SID is obtained by adding the (label_base) + (advertised SID index) = node prefix SID. The base label value for each router is chosen to be 20000, so the node prefix SID for PCC-1, for example, is 20000 + 1 = 20001.

Adjacency SIDs are generated by OSPF for each interface link, and are advertised within the extended link opaque LSA using the adjacency SID sub-TLV. The following output shows the extended link opaque LSAs of PCC-1. There are two network links, so there are two LSAs, with link state IDs of 8.0.0.2 and 8.0.0.3.

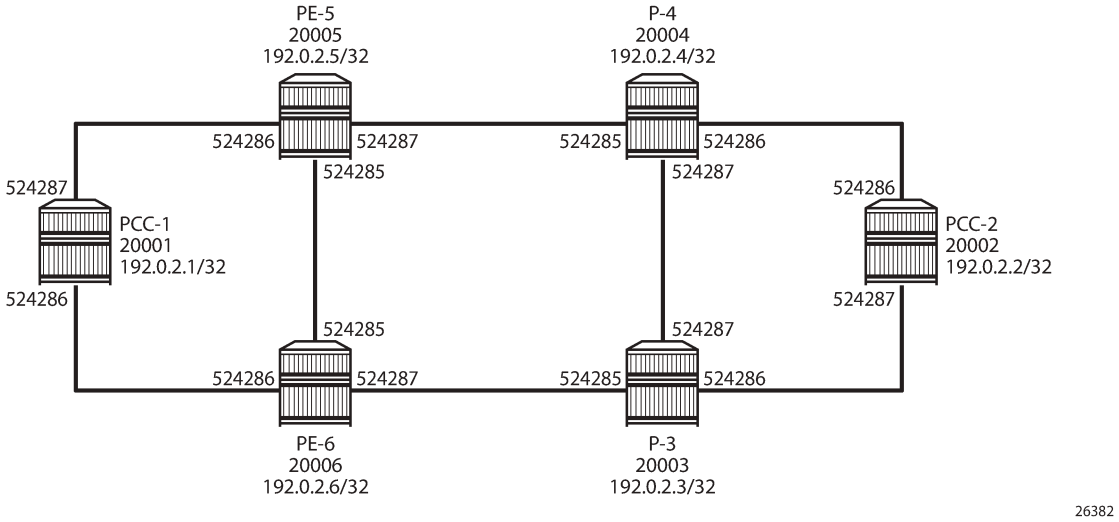
```
[/]
A:admin@PCC-1# show router ospf opaque-database adv-router 192.0.2.1 detail

===============================================================================
Rtr Base OSPFv2 Instance 0 Opaque Link State Database (type: All) (detail)
===============================================================================
---snip---

-------------------------------------------------------------------------------
Opaque LSA
-------------------------------------------------------------------------------
Area Id          : 0.0.0.0              Adv Router Id    : 192.0.2.1
Link State Id    : 8.0.0.2              LSA Type         : Area Opaque
Sequence No      : 0x80000001           Checksum         : 0x2f79
Age              : 185                  Length           : 48
Options          :  E
Advertisement    : Extended Link
    TLV Extended link (1) Len 24  :
        link Type=P2P (1)  Id=192.0.2.5 Data=192.168.15.1
        Sub-TLV Adj-SID (2) len 7 :
            Flags=Value Local (0x60)
```

```
            MT-ID=0 Weight=0 SID/Index/Label=524287
-------------------------------------------------------------------------------
Opaque LSA
-------------------------------------------------------------------------------
Area Id          : 0.0.0.0              Adv Router Id    : 192.0.2.1
Link State Id    : 8.0.0.3              LSA Type         : Area Opaque
Sequence No      : 0x80000001           Checksum         : 0x277f
Age              : 185                  Length           : 48
Options          :  E
Advertisement    : Extended Link
    TLV Extended link (1) Len 24  :
        link Type=P2P (1)  Id=192.0.2.6 Data=192.168.16.1
        Sub-TLV Adj-SID (2) len 7 :
            Flags=Value Local (0x60)
            MT-ID=0 Weight=0 SID/Index/Label=524286
===============================================================================
```

The adjacency SID for interface on PCC-1 toward PE-5 is 524287, and the adjacency SID for the interface toward PE-6 is 524286.

A full collection of SIDs for the whole network is shown in Figure 34: Node and adjacency SIDs.

Figure 34: Node and adjacency SIDs



Segment routing TE-LSPs

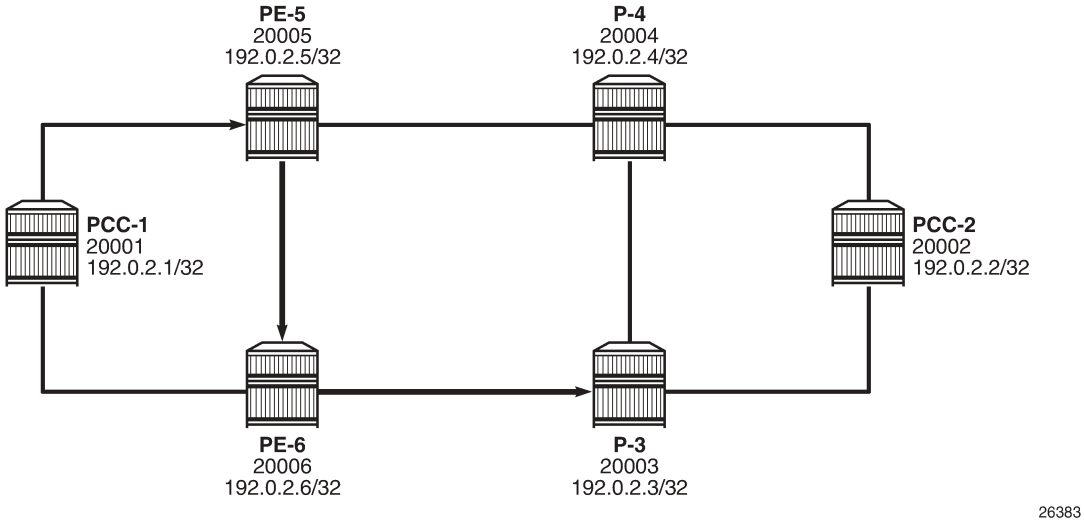
This section describes SR-TE LSPs that are configured on the head-end router (the PCC). The path taken through the network is computed locally by the PCC. To influence the path taken, a series of strict and loose hops is configured in an MPLS path.

Note:
SR-TE LSPs configured with a loose path that contains no hops is effectively a shortest path tunnel to the destination node. The destination address is resolved to the node SID of the tail-end router.

PCC-initiated and computed LSP – strict path

Consider an SR-TE LSP configured on PCC-1, with tail end at PCC-2. Assume there is a requirement for the LSP to avoid the link from PE-5 to P-4 during normal working, so a strict path from PCC-1 via PE-5 to PE-6, and then on to P-3 is required before being forwarded to PCC-2. This is shown in Figure 35: PCC computed strict path between PCC-1 and PCC-2.

Figure 35: PCC computed strict path between PCC-1 and PCC-2



To meet these requirements, an MPLS path is configured containing the following strict hops, using the system addresses to identify the hops. The following configures the MPLS path required on PCC-1. This uses the identical CLI construct as an MPLS path used in configuring an RSVP-TE LSP.

```
# on PCC-1:
configure {
    router "Base" {
        mpls {
            path "PCC-controlled-strict-path" {
                admin-state enable
                hop 1 {
                    ip-address 192.0.2.5
                    type strict
                }
                hop 2 {
                    ip-address 192.0.2.6
                    type strict
                }
                hop 3 {
                    ip-address 192.0.2.3
                    type strict
                }
            }
```

The SR-TE LSP is configured on PCC-1 as follows:

```
# on PCC-1:
configure {
```

```
    router "Base" {
        mpls {
            lsp "PCC-1-PCC-2-PCC-strict-lsp" {
                admin-state enable
                type p2p-sr-te
                to 192.0.2.2
                primary "PCC-controlled-strict-path" {
                }
            }
```

Again, the same CLI construct as an RSVP-TE LSP is used, except for the type **p2p-sr-te** . If the type is not **p2p-sr-te**, the LSP is signaled as an RSVP-TE LSP. The LSP configuration references the previously-created MPLS path as the primary path.

When **enabled**, the LSP path status is as shown in the following output:

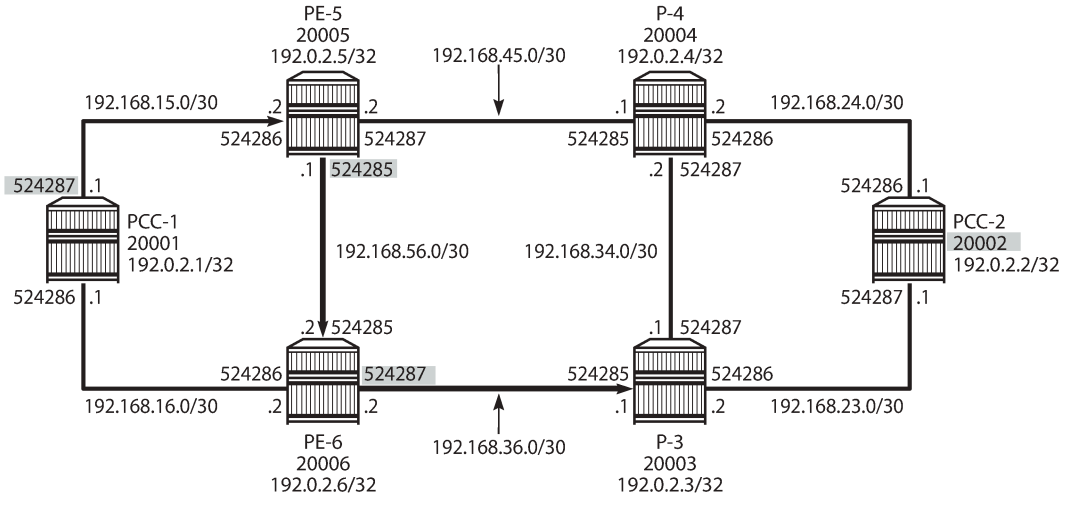
```
[/]
A:admin@PCC-1# show router mpls sr-te-lsp "PCC-1-PCC-2-PCC-strict-lsp" path detail

===============================================================================
MPLS SR-TE LSP PCC-1-PCC-2-PCC-strict-lsp
Path  (Detail)
===============================================================================
Legend :
    S      - Strict                     L      - Loose
    A-SID  - Adjacency SID              N-SID  - Node SID
    +      - Inherited
===============================================================================
-------------------------------------------------------------------------------
LSP SR-TE PCC-1-PCC-2-PCC-strict-lsp
Path  PCC-controlled-strict-path
-------------------------------------------------------------------------------
LSP Name    : PCC-1-PCC-2-PCC-strict-lsp
Path LSP ID      : 10752
From             : 192.0.2.1
To               : 192.0.2.2
Admin State      : Up                       Oper State        : Up
Path Name   : PCC-controlled-strict-path
Path Type        : Primary
Path Admin       : Up                       Path Oper         : Up
Path Up Time     : 0d 00:00:10              Path Down Time    : 0d 00:00:00
Retry Limit      : 0                        Retry Timer       : 30 sec
Retry Attempt    : 0                        Next Retry In     : 0 sec

PathCompMethod   : none                     OperPathCompMethod: none
MetricType       : igp                      Oper MetricType   : igp
LocalSrProt      : preferred                Oper LocalSrProt  : N/A
LabelStackRed    : Disabled                 Oper LabelStackRed: N/A

Bandwidth        : No Reservation           Oper Bandwidth    : 0 Mbps
Hop Limit        : 255                      Oper HopLimit     : 255
Setup Priority   : 7                        Oper SetupPriority: 7
Hold Priority    : 0                        Oper HoldPriority : 0
DelayMetricLimit : No limit                 OperDelayMetricLim: N/A
Inter-area       : N/A

PCE Updt ID      : 0                        PCE Updt State    : None
PCE Upd Fail Code: noError

PCE Report       : Disabled+                Oper PCE Report   : Disabled
PCE Control      : Disabled                 Oper PCE Control  : Disabled

Include Groups   :                          Oper IncludeGroups:
```

```
None                                            None
Exclude Groups   :                            Oper ExcludeGroups:
None                                            None
Last Resignal    : n/a

IGP/TE/Del Metric: 16777215                   Oper Metric       : 16777215
Oper MTU         : 8902                       Path Trans        : 1
Degraded         : False
Failure Code     : noError
Failure Node     : n/a
Explicit Hops    :
                    192.0.2.5(S)
                -> 192.0.2.6(S)
                -> 192.0.2.3(S)
Actual Hops      :
    192.168.15.2(192.0.2.5)(A-SID)                 Record Label        : 524287
 -> 192.168.56.2(192.0.2.6)(A-SID)                 Record Label        : 524285
 -> 192.168.36.1(192.0.2.3)(A-SID)                 Record Label        : 524287
 -> 192.0.2.2(192.0.2.2)(N-SID)                    Record Label        : 20002

BFD Configuration and State
Template         : None                       Ping Interval     : N/A
Enable           : False                      State             : notApplicable
ReturnPathLabel  : None
BFD Source Addr  : None
WaitForUpTimer   : 4 sec                      OperWaitForUpTimer: 0 sec
WaitForUpTmLeft  : 0
StartFail Rsn    : N/A

===============================================================================
```

The Actual Hops output shows the address of the upstream router facing the configured strict hop (in brackets) referenced in the MPLS path, plus a loose hop for the destination hop of 192.0.2.2.

The interface addresses are translated into SIDs to be used as MPLS labels, by the head-end PCC router, PCC-1, by examining the OSPF TE database. Each strict hop is always translated into an adjacency SID (A-SID), and a loose hop is always translated into a node SID (N-SID). This is shown in Figure 36: PCC computed LSP hop-to-label translation.

Figure 36: PCC computed LSP hop-to-label translation



When the LSP is connected, the Tunnel Table Manager (TTM) adds an entry for the SR-TE LSP. This LSP is available for the provisioning of services that use the TTM. The following output shows the tunnel table for PCC-1, which includes the shortest-path tunnels to all other routers in the network, plus the entry for the provisioned SR-TE LSP. The default preference for an SR-TE LSP in the tunnel table is 8.

```
[/]
A:admin@PCC-1# show router tunnel-table

===============================================================================
IPv4 Tunnel Table (Router: Base)
===============================================================================
Destination           Owner     Encap TunnelId  Pref   Nexthop        Metric
   Color
-------------------------------------------------------------------------------
192.0.2.2/32          sr-te     MPLS  655362    8      192.168.15.2   16777215
192.0.2.2/32          ospf (0)  MPLS  524291    10     192.168.15.2   2100
192.0.2.3/32          ospf (0)  MPLS  524294    10     192.168.16.2   1100
192.0.2.4/32          ospf (0)  MPLS  524292    10     192.168.15.2   1100
192.0.2.5/32          ospf (0)  MPLS  524293    10     192.168.15.2   1000
192.0.2.6/32          ospf (0)  MPLS  524295    10     192.168.16.2   1000
192.168.15.2/32       ospf (0)  MPLS  524289    10     192.168.15.2   0
192.168.16.2/32       ospf (0)  MPLS  524290    10     192.168.16.2   0
-------------------------------------------------------------------------------
Flags: B = BGP or MPLS backup hop available
       L = Loop-Free Alternate (LFA) hop available
       E = Inactive best-external BGP route
       k = RIB-API or Forwarding Policy backup hop
===============================================================================
```

The value of the metric is set to 16777215 (infinity – 1), because there is no CSPF and the head-end router is unaware of the full topology between head- and tail-end router.

PCC-initiated and computed LSP – loose path

Consider an LSP configured on PCC-2, with the tail end at PCC-1. There is a requirement for traffic on the LSP to pass through PE-6 before reaching PCC-1, so a loose path of PCC-2 to PE-6 before being forwarded to PCC-1 is required.

Figure 37: SR-TE LSP with loose path

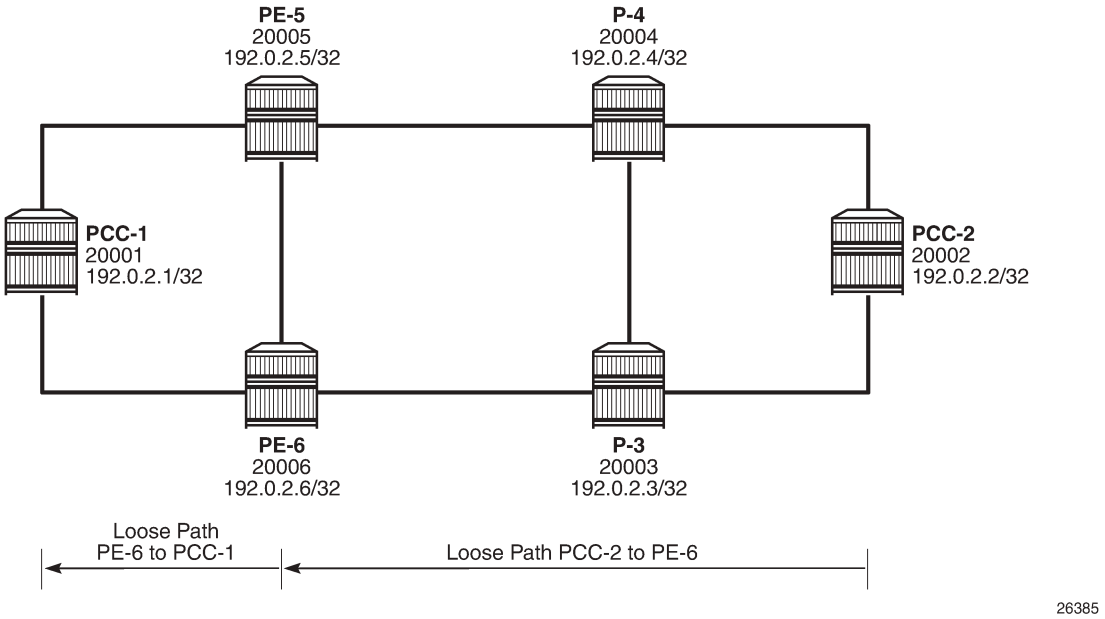


Figure 37: SR-TE LSP with loose path shows the concept of the loose path. The following configures the MPLS path containing a loose hop on PCC-2:

```
# on PCC-2:
configure {
    router "Base" {
        mpls {
            path "PCC-controlled-loose-path" {
                admin-state enable
                hop 1 {
                    ip-address 192.0.2.6
                    type loose
                }
            }
```

The SR-TE LSP configuration, which references the previously created MPLS path as the primary path, is as follows:

```
# on PCC-2:
configure {
    router "Base" {
        mpls {
            lsp "PCC-2-PCC-1-PCC-loose-lsp" {
                admin-state enable
                type p2p-sr-te
                to 192.0.2.1
                primary "PCC-controlled-loose-path" {
                }
            }
```

When enabled, the LSP path status becomes operationally up, as in the following output:

```
[/]
```

```
A:admin@PCC-2# show router mpls sr-te-lsp "PCC-2-PCC-1-PCC-loose-lsp" path detail

===============================================================================
MPLS SR-TE LSP PCC-2-PCC-1-PCC-loose-lsp
Path  (Detail)
===============================================================================
Legend :
    S      - Strict                     L      - Loose
    A-SID  - Adjacency SID              N-SID  - Node SID
    +      - Inherited
===============================================================================
-------------------------------------------------------------------------------
LSP SR-TE PCC-2-PCC-1-PCC-loose-lsp
Path  PCC-controlled-loose-path
-------------------------------------------------------------------------------
LSP Name    : PCC-2-PCC-1-PCC-loose-lsp
Path LSP ID      : 52224
From             : 192.0.2.2
To               : 192.0.2.1
Admin State      : Up                       Oper State        : Up
Path Name   : PCC-controlled-loose-path
Path Type        : Primary
Path Admin       : Up                       Path Oper         : Up
Path Up Time     : 0d 00:00:10              Path Down Time    : 0d 00:00:00
Retry Limit      : 0                        Retry Timer       : 30 sec
Retry Attempt    : 0                        Next Retry In     : 0 sec

PathCompMethod   : none                     OperPathCompMethod: none
MetricType       : igp                      Oper MetricType   : igp
LocalSrProt      : preferred                Oper LocalSrProt  : N/A
LabelStackRed    : Disabled                 Oper LabelStackRed: N/A

Bandwidth        : No Reservation           Oper Bandwidth    : 0 Mbps
Hop Limit        : 255                      Oper HopLimit     : 255
Setup Priority   : 7                        Oper SetupPriority: 7
Hold Priority    : 0                        Oper HoldPriority : 0
DelayMetricLimit : No limit                 OperDelayMetricLim: N/A
Inter-area       : N/A

PCE Updt ID      : 0                        PCE Updt State    : None
PCE Upd Fail Code: noError

PCE Report       : Disabled+                Oper PCE Report   : Disabled
PCE Control      : Disabled                 Oper PCE Control  : Disabled

Include Groups   :                          Oper IncludeGroups:
None                                            None
Exclude Groups   :                          Oper ExcludeGroups:
None                                            None
Last Resignal    : n/a

IGP/TE/Del Metric: 16777215                 Oper Metric       : 16777215
Oper MTU         : 8910                     Path Trans        : 1
Degraded         : False
Failure Code     : noError
Failure Node     : n/a
Explicit Hops    :
                   192.0.2.6(L)
Actual Hops      :
    192.0.2.6(192.0.2.6)(N-SID)                  Record Label        : 20006
 -> 192.0.2.1(192.0.2.1)(N-SID)                  Record Label        : 20001

BFD Configuration and State
Template         : None                     Ping Interval     : N/A
```

```
Enable           : False                  State            : notApplicable
ReturnPathLabel  : None
BFD Source Addr  : None
WaitForUpTimer   : 4 sec                  OperWaitForUpTimer: 0 sec
WaitForUpTmLeft  : 0
StartFail Rsn    : N/A

===============================================================================
```

The Actual Hops in the MPLS path are the configured loose hop plus a hop for the destination of 192.0.2.1.

Again, the configured hop addresses are translated into labels by the head-end PCC router, PCC-2, by examining the OSPF TE database. The hop-to-label translation always translates a loose hop to a node SID (N-SID).

The LSP is installed by the TTM into the tunnel table, alongside OSPF advertised shortest path tunnels, for use by the TTM users.

```
[/]
A:admin@PCC-2# show router tunnel-table

===============================================================================
IPv4 Tunnel Table (Router: Base)
===============================================================================
Destination           Owner     Encap TunnelId  Pref   Nexthop        Metric
   Color
-------------------------------------------------------------------------------
192.0.2.1/32          sr-te     MPLS  655362    8      192.0.2.6      16777215
192.0.2.1/32          ospf (0)  MPLS  524291    10     192.168.23.2   2100
192.0.2.3/32          ospf (0)  MPLS  524292    10     192.168.23.2   1000
192.0.2.4/32          ospf (0)  MPLS  524293    10     192.168.24.2   1000
192.0.2.5/32          ospf (0)  MPLS  524294    10     192.168.24.2   1100
192.0.2.6/32          ospf (0)  MPLS  524295    10     192.168.23.2   1100
192.168.23.2/32       ospf (0)  MPLS  524289    10     192.168.23.2   0
192.168.24.2/32       ospf (0)  MPLS  524290    10     192.168.24.2   0
-------------------------------------------------------------------------------
Flags: B = BGP or MPLS backup hop available
       L = Loop-Free Alternate (LFA) hop available
       E = Inactive best-external BGP route
       k = RIB-API or Forwarding Policy backup hop
===============================================================================
```

Service provisioning – VPRN

SR-TE tunnels are another MPLS tunnel type, and can be used in the context of **auto-bind-tunnel** for resolving BGP next hops for IPv4 routes within a VPRN.

Figure 38: VPRN service schematic

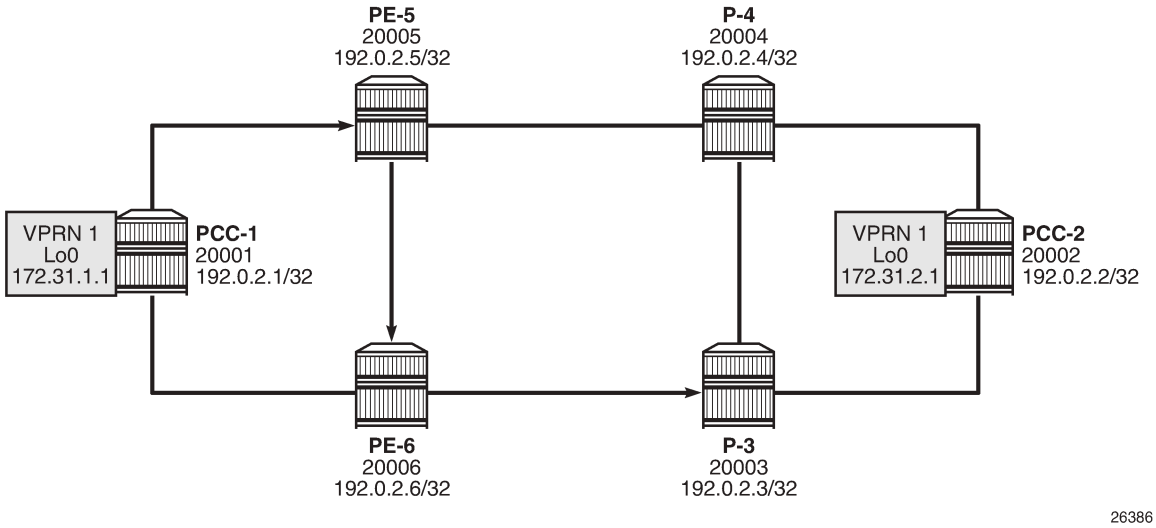


Figure 38: VPRN service schematic shows a VPRN service, configured on PCC-1 and PCC-2. The following configures the VPRN 1 on PCC-1. It includes a local interface using a /32 loopback address to be used to verify that routing is working correctly.

```
# on PCC-1:
configure {
    service {
        vprn "VPRN 1" {
            admin-state enable
            service-id 1
            customer "1"
            autonomous-system 65545
            bgp-ipvpn {
                mpls {
                    admin-state enable
                    route-distinguisher "65545:1"
                    vrf-target {
                        community "target:65545:1"
                    }
                    auto-bind-tunnel {
                        resolution filter
                        resolution-filter {
                            sr-te true
                        }
                    }
                }
            }
            interface "loopback" {
                loopback true
                ipv4 {
                    primary {
                        address 172.31.1.1
                        prefix-length 32
                    }
                }
            }
        }
```

Note:
The **auto-bind-tunnel** command has the **resolution-filter** option set to **sr-te**, so that any BGP routes received have the next-hop resolved to an SR-TE LSP. The VPRN configuration on PCC-2 also uses **auto-bind-tunnel sr-te**.

```
# on PCC-2:
configure {
    service {
        vprn "VPRN 1" {
            admin-state enable
            service-id 1
            customer "1"
            autonomous-system 65545
            bgp-ipvpn {
                mpls {
                    admin-state enable
                    route-distinguisher "65545:1"
                    vrf-target {
                        community "target:65545:1"
                    }
                    auto-bind-tunnel {
                        resolution filter
                        resolution-filter {
                            sr-te true
                        }
                    }
                }
            }
            interface "loopback" {
                loopback true
                ipv4 {
                    primary {
                        address 172.31.2.1
                        prefix-length 32
                    }
                }
            }
        }
```

Examination of the VPRN route table shows that the route prefix representing the IP address of the loopback address configured in VPRN 1 is shown, and is resolved via the SR-TE tunnel.

```
[/]
A:admin@PCC-1# show router 1 route-table

===============================================================================
Route Table (Service: 1)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
172.31.1.1/32                                 Local   Local     00h02m01s  0
       loopback                                                      0
172.31.2.1/32                                 Remote  BGP VPN   00h00m57s  170
       192.0.2.2 (tunneled:SR-TE:655362)                             16777215
-------------------------------------------------------------------------------
No. of Routes: 2
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
```

```
===============================================================================
```

Connectivity is verified by sending a ping from the loopback interface within VPRN 1 on PCC-1 to the loopback address within VPRN 1 on PCC-2, as follows:

```
[/]
A:admin@PCC-1# ping 172.31.2.1 router-instance "VPRN 1" source-address 172.31.1.1
PING 172.31.2.1 56 data bytes
64 bytes from 172.31.2.1: icmp_seq=1 ttl=64 time=4.77ms.
64 bytes from 172.31.2.1: icmp_seq=2 ttl=64 time=4.92ms.
64 bytes from 172.31.2.1: icmp_seq=3 ttl=64 time=4.91ms.
64 bytes from 172.31.2.1: icmp_seq=4 ttl=64 time=4.68ms.
64 bytes from 172.31.2.1: icmp_seq=5 ttl=64 time=4.17ms.

---- 172.31.2.1 PING Statistics ----
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min = 4.17ms, avg = 4.69ms, max = 4.92ms, stddev = 0.273ms
```

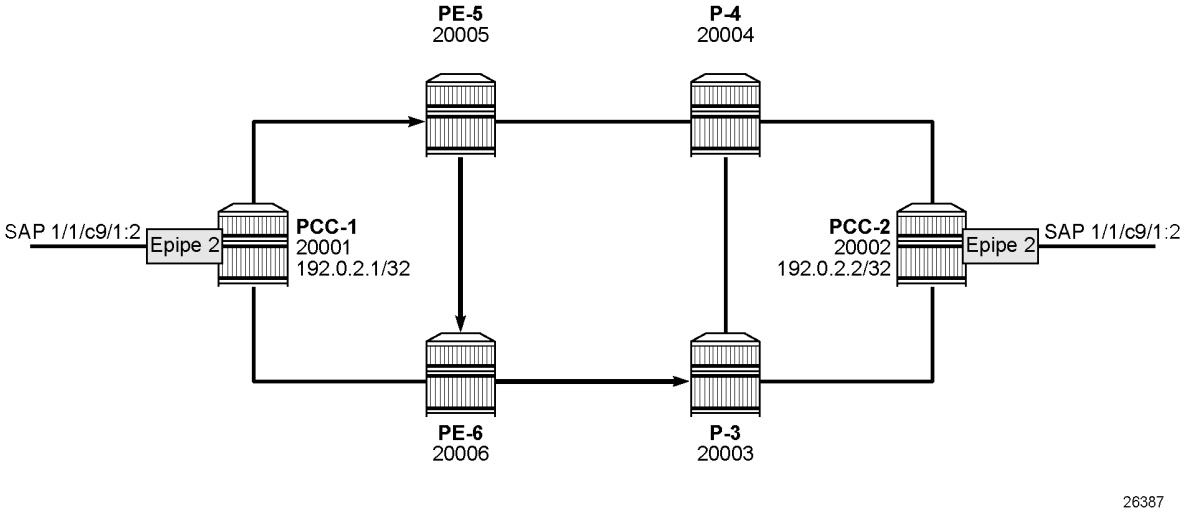
For completeness, a ping is sent in the opposite direction, between the PCC-2 VPRN 1 interface to PCC-1 VPRN 1, as follows:

```
[/]
A:admin@PCC-2# ping 172.31.1.1 router-instance "VPRN 1" source-address 172.31.2.1
PING 172.31.1.1 56 data bytes
64 bytes from 172.31.1.1: icmp_seq=1 ttl=64 time=4.98ms.
64 bytes from 172.31.1.1: icmp_seq=2 ttl=64 time=5.42ms.
64 bytes from 172.31.1.1: icmp_seq=3 ttl=64 time=5.06ms.
64 bytes from 172.31.1.1: icmp_seq=4 ttl=64 time=4.99ms.
64 bytes from 172.31.1.1: icmp_seq=5 ttl=64 time=4.79ms.

---- 172.31.1.1 PING Statistics ----
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min = 4.79ms, avg = 5.05ms, max = 5.42ms, stddev = 0.207ms
```

Layer 2 service provisioning – SR-TE

SR-TE tunnels can also be bound as a transport tunnel within SDPs. To illustrate this, consider the following example of a simple Epipe connected between PCC-1 and PCC-2, as shown in Figure 39: Epipe service schematic.

Figure 39: Epipe service schematic



Configure an SDP on PCC-1, with far end on PCC-2, and bind it to the previously configured SR-TE LSP:

```
# on PCC-1:
configure {
    service {
        sdp 12 {
            admin-state enable
            delivery-type mpls
            far-end {
                ip-address 192.0.2.2
            }
            lsp "PCC-1-PCC-2-PCC-strict-lsp" { }
        }
```

Configure an Epipe on PCC-1:

```
# on PCC-1:
configure {
    service {
        epipe "Epipe 2" {
            admin-state enable
            service-id 2
            customer "1"
            spoke-sdp 12:2 {
            }
            sap 1/2/1:2 {
            }
        }
```

Similarly, for PCC-2, configure an MPLS SDP and explicitly bind the SR-TE LSP, as follows:

```
# on PCC-2:
configure {
    service {
        sdp 21 {
            admin-state enable
            delivery-type mpls
```

```
            far-end {
                ip-address 192.0.2.1
            }
            lsp "PCC-2-PCC-1-PCC-loose-lsp" { }
        }
```

Configure Epipe 2 on PCC-2, referencing the SDP as a spoke-SDP:

```
# on PCC-2:
configure {
    service {
        epipe "2" {
            admin-state enable
            service-id 2
            customer "1"
            spoke-sdp 21:2 {
            }
            sap 1/2/1:2 {
            }
        }
```

Service verification

The state of SDP 12 on PCC-1 is shown in the following output:

```
[/]
A:admin@PCC-1# show service sdp

===============================================================================
Services: Service Destination Points
===============================================================================
SdpId  AdmMTU  OprMTU  Far End          Adm  Opr         Del     LSP   Sig
-------------------------------------------------------------------------------
12     0       8898    192.0.2.2        Up   Up          MPLS    T     TLDP
-------------------------------------------------------------------------------
Number of SDPs : 1
-------------------------------------------------------------------------------
Legend: R = RSVP, L = LDP, B = BGP, M = MPLS-TP, n/a = Not Applicable
        I = SR-ISIS, O = SR-OSPF, T = SR-TE, F = FPE
===============================================================================
```

The output shows the LSP type as an SR-TE LSP - "T".

On PCC-1, the following output shows the base state of the Epipe service entities:

```
[/]
A:admin@PCC-1# show service id 2 base

===============================================================================
Service Basic Information
===============================================================================
Service Id        : 2                   Vpn Id            : 0
Service Type      : Epipe
MACSec enabled    : no
Name              : Epipe 2
Description       : (Not Specified)
Customer Id       : 1                   Creation Origin   : manual
Last Status Change: 01/08/2026 10:15:16
Last Mgmt Change  : 01/08/2026 10:15:01
Test Service      : No
```

```
Admin State       : Up                  Oper State        : Up
MTU               : 1514
Vc Switching      : False
SAP Count         : 1                   SDP Bind Count    : 1
Per Svc Hashing   : Disabled            Lbl Eth/IP L4 TEID: Disabled
Ignore MTU Mismat*: Disabled
Vxlan Src Tep Ip  : N/A
Force QTag Fwd    : Disabled
Lcl Switch Svc St : sap
Oper Group        : <none>

-------------------------------------------------------------------------------
Service Access & Destination Points
-------------------------------------------------------------------------------
Identifier                           Type         AdmMTU  OprMTU  Adm  Opr
-------------------------------------------------------------------------------
sap:1/1/c9/1:2                        q-tag        8936    8936    Up   Up
sdp:12:2 S(192.0.2.2)                 Spok         0       8898    Up   Up
===============================================================================
* indicates that the corresponding row element may have been truncated.
```

Similarly, on PCC-2, the status of SDP 21 is as follows:

```
[/]
A:admin@PCC-2# show service sdp

===============================================================================
Services: Service Destination Points
===============================================================================
SdpId  AdmMTU  OprMTU  Far End          Adm  Opr         Del     LSP   Sig
-------------------------------------------------------------------------------
21     0       8906    192.0.2.1        Up   Up          MPLS    T     TLDP
-------------------------------------------------------------------------------
Number of SDPs : 1
-------------------------------------------------------------------------------
Legend: R = RSVP, L = LDP, B = BGP, M = MPLS-TP, n/a = Not Applicable
        I = SR-ISIS, O = SR-OSPF, T = SR-TE, F = FPE
===============================================================================
```

The state of the Epipe service on PCC-2 is shown in the following output:

```
[/]
A:admin@PCC-2# show service id 2 base

===============================================================================
Service Basic Information
===============================================================================
Service Id        : 2                   Vpn Id            : 0
Service Type      : Epipe
MACSec enabled    : no
Name              : Epipe 2
Description       : (Not Specified)
Customer Id       : 1                   Creation Origin   : manual
Last Status Change: 01/08/2026 10:15:16
Last Mgmt Change  : 01/08/2026 10:15:08
Test Service      : No
Admin State       : Up                  Oper State        : Up
MTU               : 1514
Vc Switching      : False
SAP Count         : 1                   SDP Bind Count    : 1
Per Svc Hashing   : Disabled            Lbl Eth/IP L4 TEID: Disabled
Ignore MTU Mismat*: Disabled
Vxlan Src Tep Ip  : N/A
```

```
Force QTag Fwd    : Disabled
Lcl Switch Svc St : sap
Oper Group        : <none>

-------------------------------------------------------------------------------
Service Access & Destination Points
-------------------------------------------------------------------------------
Identifier                              Type         AdmMTU  OprMTU  Adm  Opr
-------------------------------------------------------------------------------
sap:1/1/c9/1:2                          q-tag        8936    8936    Up   Up
sdp:21:2 S(192.0.2.1)                   Spok         0       8906    Up   Up
===============================================================================
* indicates that the corresponding row element may have been truncated.
```

Conclusion

Segment routing LSPs extend the use of MPLS labels into traffic engineering applications. This chapter provides the configuration for router instantiated and controlled SR-TE LSPs along with some examples of the application in a VPRN and Epipe. The chapter also shows the associated commands and outputs that can be used for verifying and troubleshooting.

Segment Routing over IPv6

This chapter provides information about Segment Routing over IPv6.

Topics in this chapter include:

Applicability

The information and MD-CLI configuration in this chapter are based on SR OS Release 21.10.R1. Segment Routing over IPv6 (SRv6) is supported on FP4-based equipment in SR OS Release 21.5.R2 and later.

Overview

Segment Routing (SR) provides control over the forwarding paths without any need for path signaling, as described in chapter Segment Routing with IS-IS Control Plane for SR over IPv4. An SR tunnel contains a list of one or more segments. Each segment is identified by a segment identifier (SID). For SR over IPv4, the SIDs are MPLS labels from a configured SR-label range.

SRv6 provides IPv6 transport with both shortest path and source routing capabilities. SRv6 is a framework for the programmability of IPv6, which utilizes the large IPv6 address space. SRv6 data path encapsulation models each SID using a 128-bit IPv6 address, with differences for shortest-path routing and source routing.

In shortest-path routing, the destination SID is encoded in the Destination Address (DA) field of the outer IPv6 header, as shown in Table 5: SRv6 shortest path routing .

Table 5: SRv6 shortest path routing

Header type	Parameter encoding
IPv6	**Next header = IP** SA = 2001:db8::2:1 **DA= 2001:db8:aaaa:101:0:1000::**
IP	Version 4, IHL=20 SA= 10.1.2.1 DA = 10.3.2.1

Header type	Parameter encoding
	Protocol = UDP ...

In source routing, the SIDs of the nodes the packet must traverse are encoded as a SID list in the Segment Routing Header (SRH). The next SID in a segment list to forward the packet to is copied from the SRH into the DA field of the outer IPv6 header. The SID in the DA field determines the termination of the current segment. At the segment endpoint node, the next header (in this case, SRH) is examined and the next active SID is copied to the DA field. Table 6: SRv6 source routing shows an example with SRv6 source routed path segment list in the SRH.

Table 6: SRv6 source routing

Header type	Parameter encoding
IPv6	**Next header = SRH** SA = 2001:db8::2:1 **DA= 2001:db8:aaaa:111:0:1000::**
SRH	Segments left = 2 Segment 0 - 2001:db8:aaaa:102:0:1000:: Segment 1 - 2001:db8:aaaa:112:0:1000:: Segment 2 - 2001:db8:aaaa:111:0:1000:: Next Header = IP

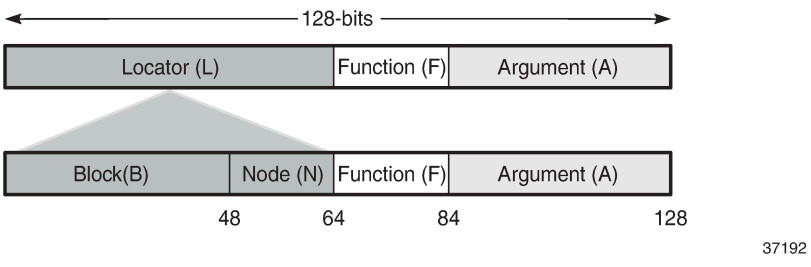
SRv6 SID

An SRv6 SID is a routable IPv6 prefix when it is set as the IPv6 header DA.

Note:
IPv6 router interface addresses are not SRv6 SIDs.

Figure 40: SRv6 SID encoding shows that the 128-bit address of an SRv6 SID is split into three constituent parts: locator, function, argument.

Figure 40: SRv6 SID encoding



- The locator is a summary IPv6 prefix for a set of SIDs instantiated on an SRv6-capable router. The locator:
  - must be explicitly configured
  - is advertised using IS-IS
  - can be associated with a topology and/or Flex-Algorithm
  - provides reachability to all SIDs originated by a router if the locator part of the SRv6 SID is routable
  - comprises the L most significant bits of the SID, with L ranging from 4 to 96 bits
  - has format B:N
    - All routers in a domain have the same block address B.
    - Each router in the domain has its own node-specific address N.
- The function is an opaque identification of a local behavior bound to the segment, as described in RFC 8986. Table 7: SRv6 endpoint behaviors supported in SR OS Release 21.10.R1 lists the SRv6 endpoint behaviors supported in SR OS Release 21.10.R1.

Table 7: SRv6 endpoint behaviors supported in SR OS Release 21.10.R1

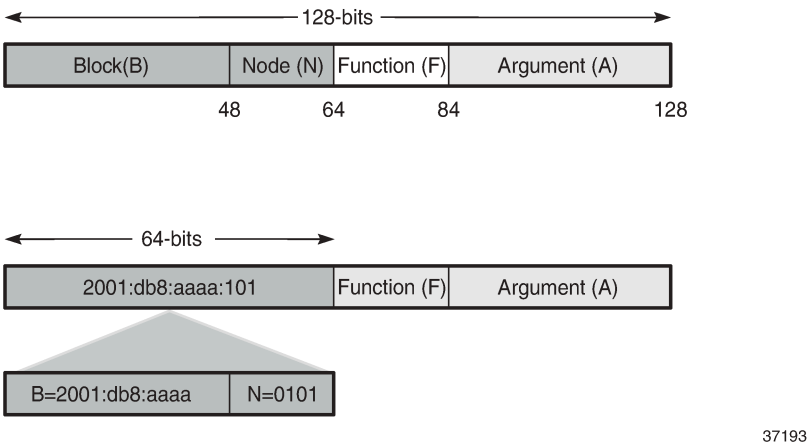
Function name	Role or behavior	Description
End	Endpoint	Equivalent to a node SID
End.X	Endpoint with an L3 cross-connect (X-connect)	Equivalent to an adjacency SID
LAN-End.X	Endpoint with an L3 cross-connect (X-connect)	Equivalent to an adjacency SID associated with a broadcast interface
End.DT4	De-encapsulate and perform an IPv4 table lookup	– VPRN table lookup: per-VRF SID for the VPN-IPv4 address family – Prefix lookup in the global IPv4 routing table
End.DT6	De-encapsulate and perform an IPv6 table lookup	– VPRN table lookup: per-VRF SID for the VPN-IPv6 address family – Prefix lookup in the global IPv6 routing table
End.DT46	De-encapsulate and perform IPv4 and IPv6 table lookups	– VPRN table lookup: both IPv4 and IPv6 - equivalent to per-VRF label – VPN-IPv4 and VPN-IPv6 routes are advertised with a single label in the same VRF

- The argument, which is not a configurable field in SR OS Release 21.10.R1, is set to all zeros.

Figure 41: SRv6 SID encoding example shows an example of an SRv6 SID with the following:

- B = 48 bits
- N = 16 bits
- L = B + N = 48 bits + 16 bits = 64 bits
- F = 20 bits
- The remaining 44 bits (A) are set to zero.

Figure 41: SRv6 SID encoding example



The /64 locator part for a set of routers in a routing domain consists of:

- a common 48-bit block, for example, 2001:db8:aaaa::/48
- a unique 16-bit node identifier allocated in the range from 0000 to ffff

Some examples:

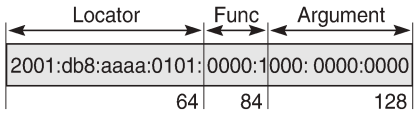
- locator for PE-1 = 2001:db8:aaaa:101::/64
- locator for PE-2 = 2001:db8:aaaa:102::/64
- locator for PE-3 = 2001:db8:aaaa:103::/64

The local router installs the locator in its IPv6 route table and FIB. The locator prefix is advertised in IS-IS in the SRv6 locator sub-TLV. Each remote router populates its route table and FIB with the locator prefixes, including the tunneled next-hop to the originating router.

The function field has a configurable length, ranging from 20 to 96 bits. By default, the function field has 20 bits. The function field is used to assign End and End.X SIDs, which are used by remote routers to create repair tunnels for remote and topology-independent loopfree-alternate (RLFA and TI-LFA) backup paths.

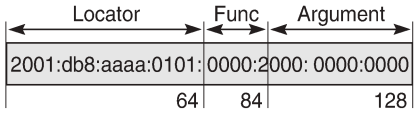
- An End function is statically configured in SR OS:
  - By default, the number of static functions is 1.
  - For example, the End function with value 1 in the 20-bit format is represented as 00001 in hexadecimal, followed by the zeros of the argument field.
  - The End SID (node SID) for PE-1 equals 2001:db8:aaaa:101:0:1000::/128, as shown in Figure 42: End SID for PE-1

Figure 42: End SID for PE-1



- The End.X function can be statically configured or automatically assigned by the system.
  - In case of static configuration, the number of static functions must be increased.
  - For the function with value 2 in a function field of 20 bits, the corresponding hexadecimal pattern is 00002, followed by the zeros of the argument field.
  - The End.X SID (adjacency SID) for PE-1 equals 2001:db8:aaaa:101:0:2000::/128, as shown in Figure 43: End.X SID for PE-1.

    Figure 43: End.X SID for PE-1



IPv6 header and SRH

This section describes how source routing works with the insertion of an SRH.

Note:
SR OS Release 21.10 has no mechanism for computing a source-routed path for normal data traffic flow; for example, there is no equivalent to the SR-TE or SR-policy label stack for source routing. The use of the SRH is restricted to repair tunnels computed by the TI-LFA process. When a link or node failure occurs, the Point of Local Repair (PLR) inserts an appropriate SRH for SRv6 traffic that is to be routed around the failure during IGP convergence.

Different SR node types are defined: source node, transit node, and segment endpoint node. To enable source routing on the IPv6 source router, the SRH contains an ordered list of one or more SRv6 SIDs.

Figure 44: IPv6 header defined in RFC 8200 shows the IPv6 header where the next header field must be coded as 43 when the IPv6 extension header, which follows the IPv6 header, is an SRH.

Figure 44: IPv6 header defined in RFC 8200

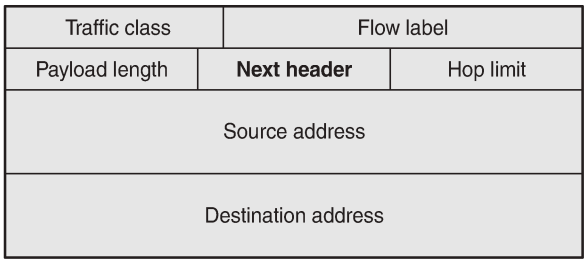


Figure 45: Position of the SRH in the protocol stack shows that the header following the IPv6 header sits between the IPv6 header and upper layer protocols, such as TCP or UDP.

Figure 45: Position of the SRH in the protocol stack

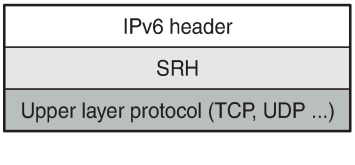
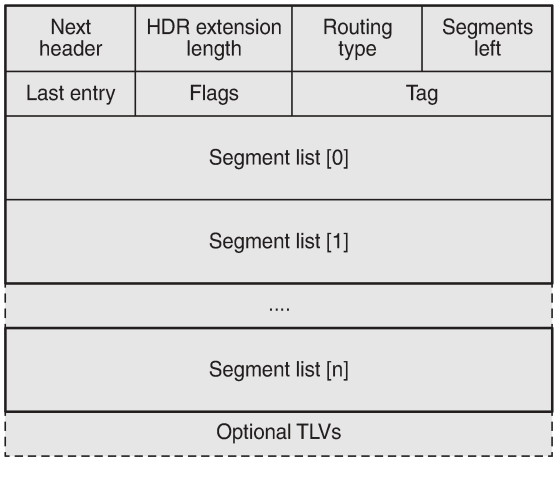


Figure 46: SRH defined in RFC 8754 shows the SRH.

Figure 46: SRH defined in RFC 8754



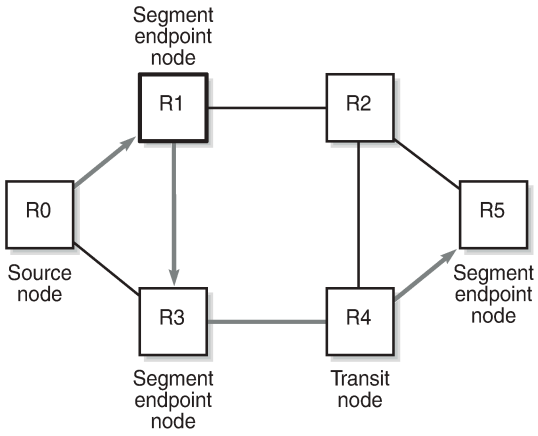
The SRH is derived from the IPv6 routing header as defined in RFC 8200. The SRH fields are:

- Next header: defines the type of header following SRH, for example, TCP or UDP.
- Routing type for SRH: 4.
- Segments left: the number of explicitly listed intermediate nodes still to be traversed before reaching the final destination.

- Last entry: contains the zero-based index of the last element of the segment list.
- Segment list [n]: a 128-bit IPv6 address representing the n^{th} segment in the segment list. The segment list is encoded in reverse numerical order: segment list [0] is the first element in the segment list and contains the last segment of the SR path, segment list [1] contains the penultimate segment of the SR path, and so on.

Figure 47: SRv6 node types shows the SR node types: source node, transit node, and segment endpoint node for an SRv6 packet flow from R0 to R5 via hops R1, R3, and R5.

Figure 47: SRv6 node types



The intermediate hops R1, R3, and R5 are programmed in the segment list of the SRH.

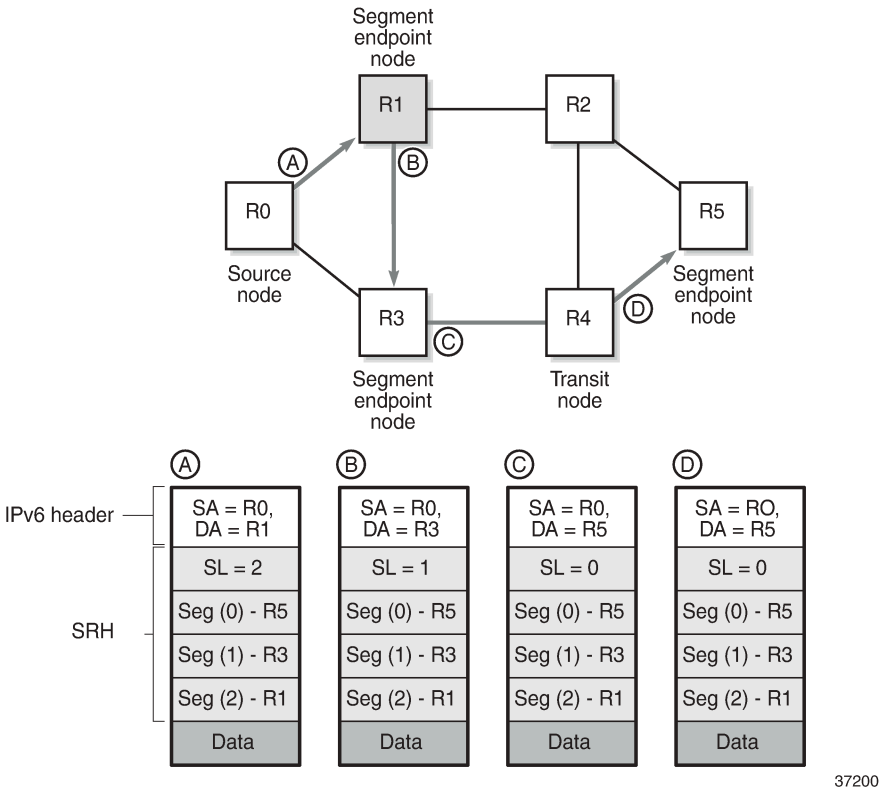
The SRv6 node types defined in RFC 8754 are:

- SR source node
  - Any node that originates an IPv6 packet with a segment (that is, an SRv6 SID) in the DA field of the IPv6 header.
  - The IPv6 packet leaving the SR source node may or may not contain an SRH. This includes either:
    - a host originating an IPv6 packet
    - an SR domain ingress router encapsulating a received packet in an outer IPv6 header, followed by an optional SRH
  - In this example, R0 acts as an SR source node and includes an SRH containing a segment list.
- SR transit node
  - Any node forwarding an IPv6 packet where the DA of the packet is not locally configured as a segment or a local interface. A transit node need not be capable of processing a segment or SRH.
  - In this example, R4 acts as an SR transit node. It forwards the SRv6 packet without processing the SRH.
- SR segment endpoint node
  - Any node receiving an IPv6 packet where the DA of that packet is locally configured as a segment or local interface.

– In this example, R1, R3, and R5 are SR segment endpoint nodes. These nodes interrogate the SRH as part of packet processing.

Figure 48: Data forwarding of SRv6 encapsulated packets using SRv6 SIDs shows the data forwarding of SRv6 encapsulated packets using SRv6 SIDs at R0 and R1.

Figure 48: Data forwarding of SRv6 encapsulated packets using SRv6 SIDs



Source node R0 tunnels an SRv6 packet to destination R5, segment (0), in the SRH.

- The segment list contains SRv6 SIDs associated with each hop, such as the End SID. The first segment endpoint is the last segment in the list, segment (2) in the example. The Segments Left (SL) field is set to a value matching the highest segment list number (2).
- SRH is only used by routers where the DA is equal to a local address. The IPv6 source address is set to the local IPv6 address of R0. The IPv6 DA in the IPv6 header is set to the segment list entry indexed in the SL field; in this case, R_1.
- The packet is forwarded to R1.

At R_1, the incoming packet has the IPv6 DA matching R1.

- R1 removes the IPv6 header and processes the SRH. The SL is decremented to SL 1, which corresponds to segment (1) = R3.
- R1 adds an IPv6 header with DA equal to the SID for R_3.
- R1 forwards the packet to R3.

At R_3, the incoming packet has the IPv6 DA matching R_3.

- R3 removes the IPv6 header and processes the SRH. The SL is decremented to SL 0, which corresponds to segment (0) = R5.
- R3 adds an IPv6 header with DA equal to the SID for R5.
- R3 forwards the packet to R5.

At R4, the incoming packet has the IPv6 DA matching R5, so the packet is forwarded to R5 without processing the SRH header and without changing the IPv6 DA.

At R5, the incoming packet has the IPv6 DA matching R_5, so the IPv6 header is removed and the SRH header is processed. The SL value 0 cannot be decreased anymore, so R5 removes the SRH and the packet is sent for further processing, for example, to a particular VPRN.

Note:
The IPv6 SID at segment (0) may contain an opaque behavior value (function) that indicates to the destination node that further processing is required, such as a VPRN table lookup.

Data path support: forwarding path extensions

SRv6 data traffic requires additional processing at both the ingress and egress data planes. This processing is performed via an internal cross connect in the form of port cross-connect (PXC) ports. SRv6 traffic is steered from the input to a PXC port, where it is internally looped for additional processing.

The PXC port is associated with the SRv6 application using a Forwarding Path Extension (FPE). An origination (egress) FPE and termination (ingress) FPE are associated with SRv6. The additional processing in the SRv6 data path is as follows:

- Ingress PE node
  - At the origination FPE data path, L2 and L3 service packets are received and the SRv6 encapsulation header is pushed for the primary path and for the backup path based on the index passed by the service context in the internal packet header.
  - The hop-limit field in the outer IPv6 header of the SRv6 tunnel is set to 255.
  - At the termination FPE data path, a lookup is done on the DA field in the outer IPv6 header, and the packet is forwarded to one of the candidate egress network IP interfaces based on a hash of the flow label and SA/DA fields of the outer IPv6 packet header.
- Egress PE node
  - At the origination FPE data path of the incoming router interface, a longest prefix match of the DA in the outer IPv6 header is performed in the SRv6 SID FIB.
  - If there is a match against a local locator prefix, the packet is forwarded to the termination FPE for service SID termination processing.
  - The termination FPE does an Ingress Label Map (ILM) lookup on the service label and forwards the packet to the service context for further processing.
  - At the origination FPE data path, the SRH is processed. The SRv6 encapsulation is removed, and a service label is inserted into the inner packet with label value derived from the function field.
- Transit router
  - Transit routers do not require FPEs.
  - Transit routers receiving SRv6-encapsulated packets make forwarding decisions based on the IPv6 route table lookups.

SRH processing modes

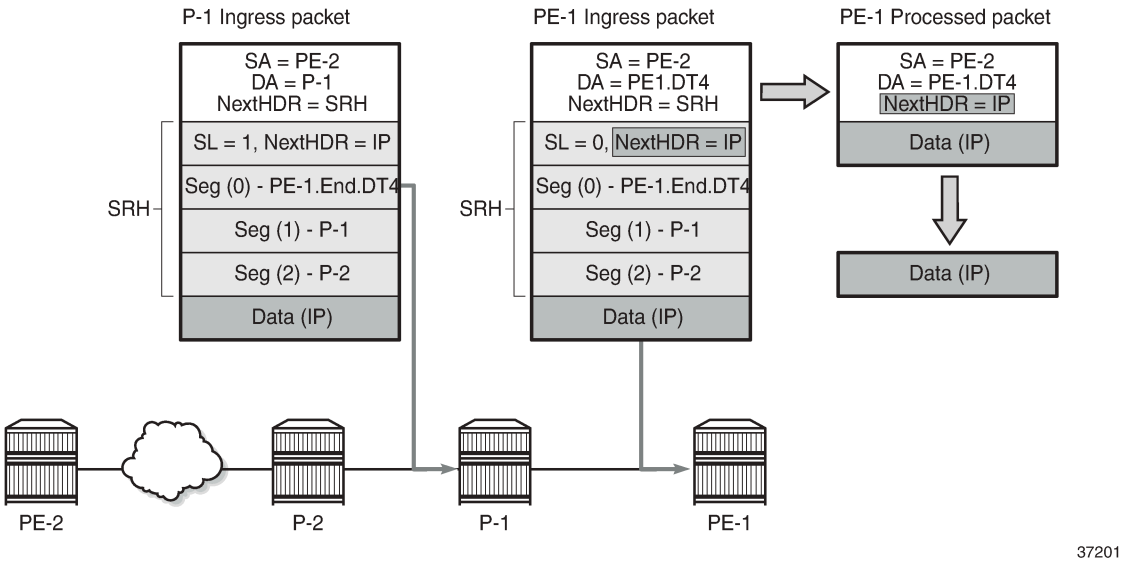
SR OS supports two SRH processing modes at the end of the SRv6 tunnel:

- Ultimate SRH Pop (USP), where the ultimate SR segment endpoint node processes and removes the SRH
- Penultimate SRH Pop (PSP), where the penultimate SR segment endpoint node processes and removes the SRH

USP mode

In the following example, source node PE-2 sends a packet to SR segment endpoint node PE-1 via intermediate hops P-2 and P-1. Figure 49: USP mode - egress router PE-1 processes and removes SRH shows how penultimate SR segment endpoint node P-1 and ultimate SR segment endpoint node PE-1 process the SRH in the packet.

Figure 49: USP mode - egress router PE-1 processes and removes SRH



Source node PE-2 sends a packet with SRH with three segments in the segment list: Seg(2) for P-2, Seg(1) for P-1, and Seg(0) for destination PE-1. P-1, P-2, and PE-1 are SR segment endpoint nodes. Penultimate SR segment endpoint node P-1 decrements the value in the SL field in the SRH from 1 to 0 and copies the PE-1 SID from segment 0 into the IPv6 header DA. Ultimate SR segment endpoint node PE-1 receives the packet with the DA equal to PE-1.End.DT4 and SL 0 and processes the packet by:

- "updating the Next Header field in the IPv6 header with the Next Header field of the SRH
- "removing the SRH from the IPv6 extension header chain
- "processing the next header in the packet, which is achieved using the origination FPE data path

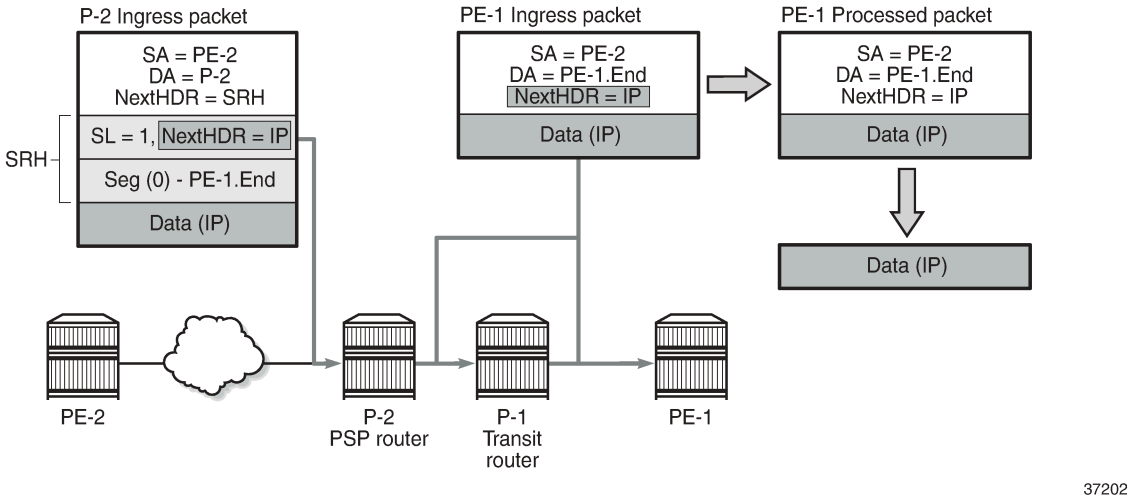
Note:
In this example, the IPv6 SID at segment (0) "PE-1.End.DT4" contains a function indicating to the destination node that a VPRN table lookup is required.

PSP mode

As stated in RFC 8986, a penultimate SR segment endpoint node is one that, as part of the SID processing, copies the last SID from the SRH into the IPv6 DA and decrements the SL value from one to zero.

Figure 50: Penultimate SRH hop P-2 processes and removes the SRH shows how penultimate SR segment endpoint node P-2 processes the packet toward PE-1. P-1 is an SR transit node in this example, so it does not process an SRH.

Figure 50: Penultimate SRH hop P-2 processes and removes the SRH



The PSP operation is controlled by the SR source node. SR source node PE-2 is aware that the PE-1.End SID has SRH mode PSP. PE-2 sends a packet to PE-1 with the DA set to P-2 in the IPv6 header. The SRH contains one SID in the segment list: Seg(0) PE-1.End. The SL is set to 1.

Penultimate SR segment endpoint node P-2 processes the packet by:

- decrementing the IPv6 hop limit by 1
- decrementing the SL by 1, so SL = 0
- updating the IPv6 DA with the PE-1.End node SID from the segment list
- updating the Next Header field in the IPv6 header to the Next Header field of the SRH
- removing the SRH from the IPv6 extension header chain
- submitting the packet to the MPLS engine for transmission

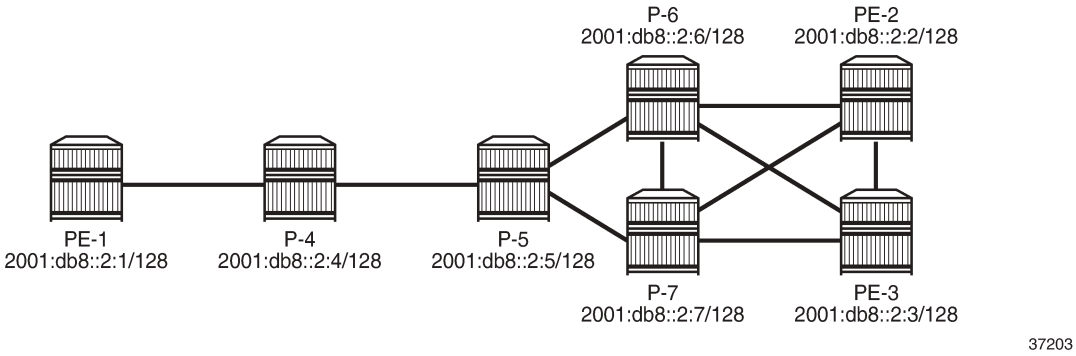
At transit node P-1, the packet is forwarded based on the RTM lookup for IPv6 DA in the IPv6 header.

At the destination node PE-1, the IPv6 header is removed and additional processing of the next header in the packet is done via an origination FPE data path.

Configuration

Figure 51: Example topology shows the example topology with seven SRv6-capable routers (with FP4).

Figure 51: Example topology



PXC

SRv6 traffic is steered from input to a PXC port, where it is internally looped for additional processing. PXC can use either an internally looped physical port or an internal loopback in the FP4 MAC chip.

In case of an internally looped physical port, configure PXC on the physical port, as shown in the following example for PXC 5 on physical port 1/1/c5/1:

```
# on all SRv6-capable nodes:
configure {
    port-xc {
        pxc 5 {
            admin-state enable
            port-id 1/1/c5/1
        }
    }
    port pxc-5.a {
        admin-state enable
    }
    port pxc-5.b {
        admin-state enable
    }
    port 1/1/c5 {
        admin-state enable
        connector {
            breakout c4-10g
        }
    }
    port 1/1/c5/1 {
        admin-state enable
        ethernet {
            mode hybrid
            dot1x {
                tunneling true
            }
        }
```

```
    }
```

In case of internal loopbacks in the FP4 MAC chip, map PXC 1 and PXC 2 to internal loopbacks. It is possible to map PXC 1 and PXC 2 to the same loopback on the same MAC chip, but that is not configured here.

```
# on all SRv6-capable nodes:
configure {
    card 1 {
        card-type xcm-2s
        mda 1 {
            mda-type s36-100gb-qsfp28
            xconnect {
                mac 1 {
                    loopback 1 {
                    }
                    loopback 2 {
                    }
                }
            }
        }
    }
    port-xc {
        pxc 1 {
            admin-state enable
            port-id 1/1/m1/1
        }
        pxc 2 {
            admin-state enable
            port-id 1/1/m1/2                # or loopback 1/1/m1/1 (same as PXC 1)
        }
    }
    port pxc-1.a {
        admin-state enable
    }
    port pxc-1.b {
        admin-state enable
    }
    port pxc-2.a {
        admin-state enable
    }
    port pxc-2.b {
        admin-state enable
    }
    port 1/1/m1/1 {
        admin-state enable
    }
    port 1/1/m1/2 {
        admin-state enable
    }
```

There are several MAC chips per FP4-complex (hardware dependent). The operator configures the location of the loopback. The PXC loopback must be referenced as a port ID to enable loopback. The following **show datapath** command includes the internal loopbacks 1/1/m1/1 and 1/1/m1/2:

```
[/]
A:admin@PE-2# show datapath 1 detail

===============================================================================
Card    [XIOM/]MDA  FP  TAP   MAC Chip Num   Connector      Port
-------------------------------------------------------------------------------
1       1           1   1     1             c1             1/1/c1/1
```

```
1       1           1   1     1             c1            1/1/c1/2
1       1           1   1     1             c1            1/1/c1/3
1       1           1   1     1             c1            1/1/c1/4
---snip---

1       1           2   1     6             c36           1/1/c36/1
1       1           2   1     6             c36           1/1/c36/2
1       1           2   1     6             c36           1/1/c36/3
1       1           2   1     6             c36           1/1/c36/4
1       1           1   1     1             N/A           1/1/m1/1
1       1           1   1     1             N/A           1/1/m1/2
===============================================================================
```

In this example, PXC loopbacks are configured on MDA 1/1, which has two MAC chips with MAC chip numbers m1 and m2. The two internal loopbacks are configured on MAC chip number m1.

Note:
Nokia recommends selecting cards and MAC chips connected to faceplate ports with lower bandwidth utilization for internal PXCs.

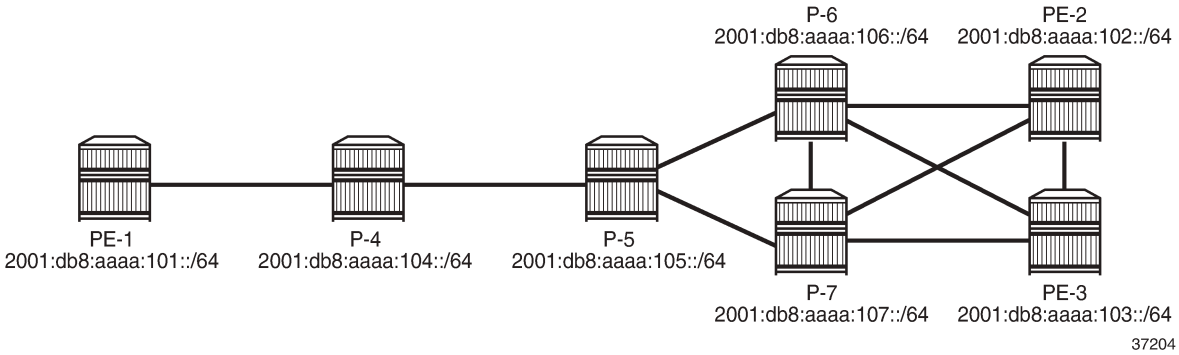
In this example, all nodes can act as SR segment endpoint nodes; there are no SR transit nodes. Perform the following steps to enable SRv6 on the nodes:

1. Allocate an address block B for all routers in a domain; for example, 2001:db8:aaaa::/48.
2. Allocate a unique node address N for each router; for example, 0101 for PE-1.
3. Configure a locator for each router in the format B:N:: and set the prefix length of the locator; for example, /64.
4. Add FPE to configure the data path.
5. Configure the End function (the SRv6 equivalent for node SID) for each router locator.
6. Configure the End.X functions (the SRv6 equivalent for adjacency SIDs) for each router associated with locator.
7. Advertise the locator in IS-IS level 1 or 2, as required.

Locator B:N::

Figure 52: SRv6 router locator prefixes shows the router locator prefixes for the seven nodes in the sample topology.

Figure 52: SRv6 router locator prefixes



Configure the SRv6 address block B and the locator prefix on the nodes. The following example shows SRv6 address block B 2001:db8:aaaa::/48 and locator prefix 2001:db8:aaaa:101::/64 in the dedicated **segment-routing-v6** context on PE-1:

```
# on PE-1:
configure {
    router "Base" {
        segment-routing {
            segment-routing-v6 {
                locator "PE-1_loc" {
                    admin-state enable
                    block-length 48
                    prefix {
                        ip-prefix 2001:db8:aaaa:101::/64
                    }
                }
            }
```

The configuration on the other nodes is similar with the locator prefixes as shown in Figure 52: SRv6 router locator prefixes.

FPE

SRv6 packet processing requires an ingress (termination) FPE and an egress (origination) FPE. FPE 1 is configured as **srv6>type origination**; FPE 2 as **srv6>type termination**. FPE 1 is configured as **origination-fpe** in the global **segment-routing-v6** context; FPE 2 is configured as **termination-fpe** in the **locator** context. On PE-1, the configuration is as follows:

```
# on PE-1:
configure {
    fwd-path-ext {
        fpe 1 {
            path {
                pxc 1
            }
            application {
                srv6 {
                    type origination
                }
            }
        }
```

```
        fpe 2 {
            path {
                pxc 2
            }
            application {
                srv6 {
                    type termination
                }
            }
        }
    }
    router "Base" {
        segment-routing {
            segment-routing-v6 {
                origination-fpe [1]
                source-address 2001:db8::2:1
                locator "PE-1_loc" {
                    admin-state enable
                    block-length 48
                    termination-fpe [2]
                    prefix {
                        ip-prefix 2001:db8:aaaa:101::/64
                    }
                }
            }
```

The configuration on the other nodes is similar.

The following command for FPE 1 shows that SRv6 is enabled and operationally up and the SRv6 type is origination:

```
[/]
A:admin@PE-1# show fwd-path-ext fpe 1

===============================================================================
FPE Id: 1
===============================================================================
Description        : (Not Specified)
Multi-Path         : Disabled
Path               : pxc 1
Pw Port            : Disabled                          Oper     : down
Sub Mgmt Extension : Disabled                          Oper     : N/A
Vxlan Termination  : Disabled                          Oper     : down
Segment-Routing V6 : Enabled                           Oper     : up
SRv6 Type          : origination
If-A Qos Policy    : default
If-B MTU           : 9786 bytes                        Oper MTU : 1556 bytes
If-B Qos Policy    : default
===============================================================================
```

The following command for FPE 2 shows that SRv6 is enabled and operationally up and the SRv6 type is termination:

```
[/]
A:admin@PE-1# show fwd-path-ext fpe 2

===============================================================================
FPE Id: 2
===============================================================================
Description        : (Not Specified)
Multi-Path         : Disabled
Path               : pxc 2
```

```
Pw Port            : Disabled                           Oper     : down
Sub Mgmt Extension : Disabled                           Oper     : N/A
Vxlan Termination  : Disabled                           Oper     : down
Segment-Routing V6 : Enabled                            Oper     : up
SRv6 Type          : termination
If-A Qos Policy    : default
If-B MTU           : 0 bytes                            Oper MTU : 1556 bytes
If-B Qos Policy    : default
===============================================================================
```

The following command on PE-1 shows the associations for FPE 1. FPE 1 is an origination FPE, so it is not associated with a locator.

```
[/]
A:admin@PE-1# show fwd-path-ext associations fpe 1

===============================================================================
Segment-routing V6 associations
===============================================================================
Srv6
-------------------------------------------------------------------------------
Origination-fpe
===============================================================================

===============================================================================
Segment-routing V6 Locator associations
===============================================================================
Locator
-------------------------------------------------------------------------------
===============================================================================
```

The following command on PE-1 shows the associations for FPE 2. FPE 2 is a termination FPE associated with locator "PE-1_loc".

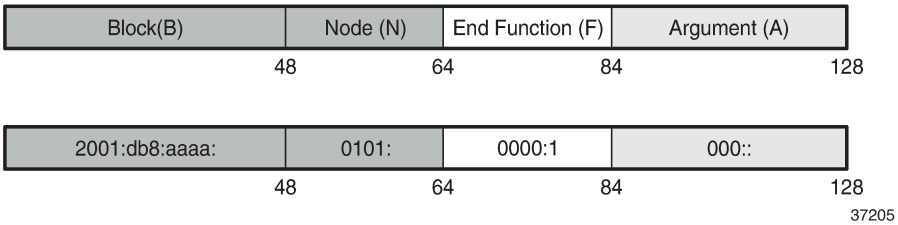
```
[/]
A:admin@PE-1# show fwd-path-ext associations fpe 2

===============================================================================
Segment-routing V6 associations
===============================================================================
Srv6
-------------------------------------------------------------------------------
===============================================================================

===============================================================================
Segment-routing V6 Locator associations
===============================================================================
Locator
-------------------------------------------------------------------------------
PE-1_loc
===============================================================================
```

Functions

End function

The SRv6 End function is configured in the SRv6 End SID that is the equivalent for IPv4 node SIDs. Figure 53: SRv6 End SID on PE-1 shows an example with End function value 1 on PE-1:

Figure 53: SRv6 End SID on PE-1



The SRv6 End function is statically configured in the **segment-routing-v6>base-routing-instance>locator** context, as follows:

```
# on PE-1:
configure {
    router "Base" {
        segment-routing {
            segment-routing-v6 {
                locator "PE-1_loc" {
                    admin-state enable
                    block-length 48
                    # function-length 20         # default value 20
                }
                base-routing-instance {
                    locator "PE-1_loc" {
                        function {
                            end 1 {              # function value = 1
                                srh-mode usp     # Ultimate SRH Pop (default: PSP)
                            }
                        }
                    }
                }
            }
```

The configuration on the other nodes is similar.

By default, the **function-length** is 20. The value **function>end 1** defines a function value of 1 inserted into the 20-bit function field. The **srh-mode** determines whether USP or PSP mode is used to process and remove the SRH. The default SRH mode is PSP.

A node can have one or two End functions. If both PSP and USP modes are used, a unique End function can be configured for each SRH mode. This requires the increase of the number static functions allowed, because the default value is 1; **static-function>max-entries** is configured for this purpose. As an example, this is configured on PE-1 only:

```
# on PE-1:
configure {
```

```
    router "Base" {
        segment-routing {
            segment-routing-v6 {
                locator "PE-1_loc" {
                    static-function {
                        max-entries 2            # 2 static functions (end 1, end 2)
                    }
                }
                base-routing-instance {
                    locator "PE-1_loc" {
                        function {
                            end 1 {              # function value = 1
                                srh-mode usp     # SRH mode - Ultimate SRH Pop
                            }
                            end 2 {              # function value = 2
                            }              # default SRH mode - Penultimate SRH Pop
                        }
                    }
                }
            }
```

The following command shows the End SID values for the locators in the base routing instance on PE-1:

```
[/]
A:admin@PE-1# show router segment-routing-v6 local-sid end

===============================================================================
Segment Routing v6 Local SIDs
===============================================================================
SID                                             Type           Function
  Locator
  Context
-------------------------------------------------------------------------------
2001:db8:aaaa:101:0:1000::                      End            1
  PE-1_loc
  Base
2001:db8:aaaa:101:0:2000::                      End            2
  PE-1_loc
  Base
-------------------------------------------------------------------------------
SIDs : 2
-------------------------------------------------------------------------------
===============================================================================
```

The following command on PE-1 shows the End SIDs plus SRH mode for the locators in the base routing instance:

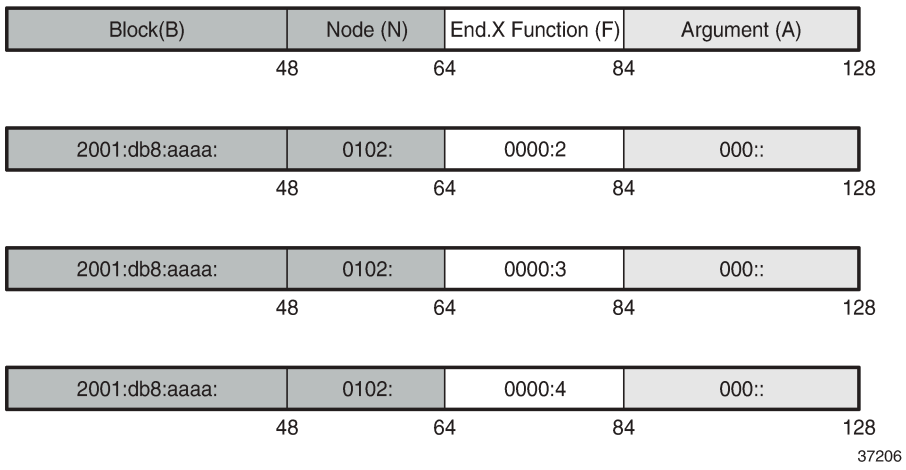
```
[/]
A:admin@PE-1# show router segment-routing-v6 base-routing-instance end

===============================================================================
Segment Routing v6 Base Routing Instance
===============================================================================
Locator
  Type       Function    SID                                     Status/InstId
    SRH-mode Protection  Interface
-------------------------------------------------------------------------------
PE-1_loc
  End                  1 2001:db8:aaaa:101:0:1000::              ok
    USP
  End                  2 2001:db8:aaaa:101:0:2000::              ok
    PSP
```

```
-------------------------------------------------------------------------------
  Auto-allocated End.X:
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
===============================================================================
Legend: * - System allocated
```

End.X function

The SRv6 End.X SID is the equivalent to IPv4 adjacency SIDs. Figure 54: SRv6 End.X SIDs on PE-2 shows an example with End.X function values 2, 3, and 4 on PE-2:

Figure 54: SRv6 End.X SIDs on PE-2



End.X function SIDs can be allocated dynamically or configured as static SIDs.

For dynamically-allocated End.X function SIDs, the configuration is as follows:

```
# on PE-2:
configure {
    router "Base" {
        segment-routing {
            segment-routing-v6 {
                base-routing-instance {
                    base-routing-instance {
                    locator "PE-2_loc" {
                        function {
                            end-x-auto-allocate usp protection protected { }
                        }
                    }
                }
            }
        }
        isis 0 {
            advertise-router-capability area
            segment-routing-v6 {
                admin-state enable
                locator "PE-2_loc" {
                    level-capability 2
                    level 2 {
                        metric 10
```

```
                    }
                }
            }
            ---snip---
```

PE-2 has three neighbors, so three End.X functions are automatically allocated. Each End.X SID is associated with a locator. In this example, the number of static functions is 1 and the automatically allocated End.X functions get values 2, 3, and 4. The PSP and USP protection modes specify whether the link is eligible for xLFA protection.

Note:
If TI-LFA is enabled, the protection mode must be set to Protected for the IGP to generate an End.X SID.

The following commands on PE-2 shows all local SIDs, including the End SID as well as End.X SIDs.

```
[/]
A:admin@PE-2# show router segment-routing-v6 local-sid

===============================================================================
Segment Routing v6 Local SIDs
===============================================================================
SID                                                 Type           Function
  Locator
  Context
-------------------------------------------------------------------------------
2001:db8:aaaa:102:0:1000::                          End            1
  PE-2_loc
  Base
2001:db8:aaaa:102:0:2000::                          End.X          2
  PE-2_loc
    None
2001:db8:aaaa:102:0:3000::                          End.X          3
  PE-2_loc
    None
2001:db8:aaaa:102:0:4000::                          End.X          4
  PE-2_loc
    None
-------------------------------------------------------------------------------
SIDs : 4
-------------------------------------------------------------------------------
===============================================================================
```

```
[/]
A:admin@PE-2# show router segment-routing-v6 base-routing-instance

===============================================================================
Segment Routing v6 Base Routing Instance
===============================================================================
Locator
  Type       Function    SID                                     Status/InstId
    SRH-mode Protection  Interface
-------------------------------------------------------------------------------
PE-2_loc
  End                  1 2001:db8:aaaa:102:0:1000::              ok
    USP
-------------------------------------------------------------------------------
  Auto-allocated End.X: USP Protected,
-------------------------------------------------------------------------------
  End.X               *2 2001:db8:aaaa:102:0:2000::              0
    USP        Protected int-PE-2-PE-3
```

```
    ISIS Level: L2 Mac Address: 02:18:01:01:00:0b Nbr Sys Id: 1920.0000.2003
  End.X               *3 2001:db8:aaaa:102:0:3000::              0
    USP        Protected int-PE-2-P-6
    ISIS Level: L2 Mac Address: 02:24:01:01:00:01 Nbr Sys Id: 1920.0000.2006
  End.X               *4 2001:db8:aaaa:102:0:4000::              0
    USP        Protected int-PE-2-P-7
    ISIS Level: L2 Mac Address: 02:28:01:01:00:15 Nbr Sys Id: 1920.0000.2007
-------------------------------------------------------------------------------
===============================================================================
Legend: * - System allocated
```

The End.X function can be created as a static SID, persistent through a reboot or link flap. The maximum number of static functions must be increased because additional static entries are required: one for each neighbor of PE-3, as follows:

```
# on PE-3:
configure {
    router "Base" {
        segment-routing {
            segment-routing-v6 {
                locator "PE-3_loc" {
                    static-function {
                        max-entries 4          # 1 End function + 3 End.X functions
                    }
                }
                base-routing-instance {
                    locator "PE-3_loc" {
                        function {
                            end 1 {
                                srh-mode usp
                            }
                            end-x 2 {
                                srh-mode usp
                                interface-name "int-PE-3-PE-2"
                            }
                            end-x 3 {
                                srh-mode usp
                                interface-name "int-PE-3-P-6"
                            }
                            end-x 4 {
                                srh-mode usp
                                interface-name "int-PE-3-P-7"
                            }
                        }
                    }
                }
            }
```

The following commands show the configured End.X SIDs on PE-3:

```
[/]
A:admin@PE-3# show router segment-routing-v6 local-sid end-x

===============================================================================
Segment Routing v6 Local SIDs
===============================================================================
SID                                                 Type           Function
  Locator
  Context
-------------------------------------------------------------------------------
2001:db8:aaaa:103:0:2000::                          End.X          2
  PE-3_loc
```

```
  Base
2001:db8:aaaa:103:0:3000::                          End.X          3
  PE-3_loc
  Base
2001:db8:aaaa:103:0:4000::                          End.X          4
  PE-3_loc
  Base
-------------------------------------------------------------------------------
SIDs : 3
-------------------------------------------------------------------------------
===============================================================================
```

```
[/]
A:admin@PE-3# show router segment-routing-v6 base-routing-instance end-x

===============================================================================
Segment Routing v6 Base Routing Instance
===============================================================================
Locator
  Type       Function    SID                                      Status/InstId
    SRH-mode Protection  Interface
-------------------------------------------------------------------------------
PE-3_loc
  End.X               2 2001:db8:aaaa:103:0:2000::               ok
    USP        Protected int-PE-3-PE-2
  End.X               3 2001:db8:aaaa:103:0:3000::               ok
    USP        Protected int-PE-3-P-6
  End.X               4 2001:db8:aaaa:103:0:4000::               ok
    USP        Protected int-PE-3-P-7
-------------------------------------------------------------------------------
  Auto-allocated End.X:
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
===============================================================================
```

SRv6 configuration summary example

The following summarizes the SRv6 configuration on PE-2:

```
# on PE-2:
configure {
    card 1 {
        card-type xcm-2s
        mda 1 {
            mda-type s36-100gb-qsfp28
            xconnect {
                mac 1 {
                    loopback 1 {          # create internal MAC-chip loopback
                    }
                    loopback 2 {
                    }
                }
            }
        }
    }
    port pxc-1.a {
        admin-state enable
    }
    port pxc-1.b {
        admin-state enable
```

```
    }
    port pxc-2.a {
        admin-state enable
    }
    port pxc-2.b {
        admin-state enable
    }
    port 1/1/m1/1 {
        admin-state enable                      # enable internal loopback port
    }
    port 1/1/m1/2 {
        admin-state enable                      # enable internal loopback port
    }
    port-xc {
        pxc 1 {
            admin-state enable
            port-id 1/1/m1/1
        }
        pxc 2 {
            admin-state enable
            port-id 1/1/m1/2
        }
    }
    fwd-path-ext {
        fpe 1 {
            path {
                pxc 1                       # map FPE 1 to PXC 1
            }
            application {
                srv6 {
                    type origination
                }
            }
        }
        fpe 2 {
            path {
                pxc 2                       # map FPE 2 to PXC 2
            }
            application {
                srv6 {
                    type termination
                }
            }
        }
    }
    router "Base" {
        segment-routing {
            segment-routing-v6 {
                origination-fpe [1]
                source-address 2001:db8::2:2
                locator "PE-2_loc" {
                    admin-state enable
                    block-length 48
                    function-length 20
                    termination-fpe [2]
                    prefix {
                        ip-prefix 2001:db8:aaaa:102::/64
                    }
                }
                base-routing-instance {
                    locator "PE-2_loc" {
                        function {
                            end 1 {
                                srh-mode usp
```

```
                            }
                            end-x-auto-allocate usp protection protected { }
                        }
                    }
                }
            }
        }
        isis 0 {
            admin-state enable
            advertise-passive-only true
            advertise-router-capability area
            ipv6-routing native
            level-capability 2
            router-id 192.0.2.2
            traffic-engineering true
            area-address [49.0001]
            loopfree-alternate {
                remote-lfa {
                }
                ti-lfa {
                }
            }
            traffic-engineering-options {
                ipv6 true
                application-link-attributes {
                }
            }
            segment-routing-v6 {
                admin-state enable
                locator "PE-2_loc" {
                    level-capability 2
                    level 2 {
                        metric 10
                    }
                }
            }
---snip---
```

Route table and tunnel table support

Each SRv6-enabled router advertises a locator prefix. Each router in the SRv6 domain installs resolved locator prefixes from received SRv6 locator TLVs. The following shows the SRv6 locator TLV on PE-2:

```
[/]
A:admin@PE-2# show router isis database PE-2 detail | match "SRv6 Locator" post-lines 5
  SRv6 Locator  :
    MT ID : 0
    Metric: ( ) 10 Algo:0
    Prefix   : 2001:db8:aaaa:102::/64
    Sub TLV   :
      End-SID   : 2001:db8:aaaa:102:0:1000::, flags:0x0, endpoint:End-USP
```

All SRv6 locators are populated in an IPv6 tunnel table and are programmed into an IPv6 FIB, so they can be displayed in the IPv6 route table. In SR OS Release 21.10 and later, a tunnel table entry is created for remote locator prefixes that have two or more ECMP next hops. A tunnel table entry is also created for remote locator prefixes with a primary and backup LFA next hop. If the remote locator prefix has no alternative path (ECMP or LFA), no tunnel table entry is created.

The IPv6 tunnel table is populated by the SR module after receipt of:

- an End or End.X SID from IS-IS
- an End.DT4 or End.DT6 SID from BGP

The following command shows the SRv6 tunnels on PE-2:

```
[/]
A:admin@PE-2# show router tunnel-table ipv6

===============================================================================
IPv6 Tunnel Table (Router: Base)
===============================================================================
Destination                                 Owner     Encap TunnelId  Pref
Nexthop                                     Color           Metric
-------------------------------------------------------------------------------
2001:db8:aaaa:101::/64 [L]                  srv6-isis SRV6  524292    0
  fe80::24:1ff:fe01:1-"int-PE-2-P-6"                          50
2001:db8:aaaa:102:0:2000::/128 [L]          srv6-isis SRV6  524289    0
  fe80::24:1ff:fe01:1-"int-PE-2-P-6"                          10
2001:db8:aaaa:102:0:3000::/128 [L]          srv6-isis SRV6  524290    0
  fe80::28:1ff:fe01:15-"int-PE-2-P-7"                         10
2001:db8:aaaa:102:0:4000::/128 [L]          srv6-isis SRV6  524291    0
  fe80::18:1ff:fe01:b-"int-PE-2-PE-3"                         10
2001:db8:aaaa:103::/64 [L]                  srv6-isis SRV6  524293    0
  fe80::18:1ff:fe01:b-"int-PE-2-PE-3"                         20
2001:db8:aaaa:104::/64 [L]                  srv6-isis SRV6  524294    0
  fe80::24:1ff:fe01:1-"int-PE-2-P-6"                          40
2001:db8:aaaa:105::/64 [L]                  srv6-isis SRV6  524295    0
  fe80::24:1ff:fe01:1-"int-PE-2-P-6"                          30
2001:db8:aaaa:106::/64 [L]                  srv6-isis SRV6  524296    0
  fe80::24:1ff:fe01:1-"int-PE-2-P-6"                          20
2001:db8:aaaa:107::/64 [L]                  srv6-isis SRV6  524297    0
  fe80::28:1ff:fe01:15-"int-PE-2-P-7"                         20
-------------------------------------------------------------------------------
Flags: B = BGP or MPLS backup hop available
       L = Loop-Free Alternate (LFA) hop available
       E = Inactive best-external BGP route
       k = RIB-API or Forwarding Policy backup hop
===============================================================================
```

The following shows the IPv6 FP-tunnel table on PE-2. For locator prefix 2001:db8:aaaa:101::/64, tunnel ID 524292 has primary next hop fe80::23:ffff:fe00:0-"int-PE-2-P-6" and backup next hop fe80::27:ffff:fe00:0-"int-PE-2-P-7".

```
[/]
A:admin@PE-2# show router fp-tunnel-table 1 ipv6

===============================================================================
IPv6 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                 Protocol         Tunnel-ID
  Lbl/SID
    NextHop                                                  Intf/Tunnel
  Lbl/SID (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
2001:db8:aaaa:101::/64                      SRV6              524292
  -
    fe80::24:1ff:fe01:1-"int-PE-2-P-6"                        1/1/c2/1:1000
```

```
  -
    fe80::28:1ff:fe01:15-"int-PE-2-P-7"(B)                    1/1/c3/1:1000
2001:db8:aaaa:103::/64                        SRV6            524293
  -
    fe80::18:1ff:fe01:b-"int-PE-2-PE-3"                       1/1/c1/1:1000
  -
    fe80::24:1ff:fe01:1-"int-PE-2-P-6"(B)                     1/1/c2/1:1000
2001:db8:aaaa:104::/64                        SRV6            524294
  -
    fe80::24:1ff:fe01:1-"int-PE-2-P-6"                        1/1/c2/1:1000
  -
    fe80::28:1ff:fe01:15-"int-PE-2-P-7"(B)                    1/1/c3/1:1000
2001:db8:aaaa:105::/64                        SRV6            524295
  -
    fe80::24:1ff:fe01:1-"int-PE-2-P-6"                        1/1/c2/1:1000
  -
    fe80::28:1ff:fe01:15-"int-PE-2-P-7"(B)                    1/1/c3/1:1000
2001:db8:aaaa:106::/64                        SRV6            524296
  -
    fe80::24:1ff:fe01:1-"int-PE-2-P-6"                        1/1/c2/1:1000
  -
    fe80::18:1ff:fe01:b-"int-PE-2-PE-3"(B)                    1/1/c1/1:1000
2001:db8:aaaa:107::/64                        SRV6            524297
  -
    fe80::28:1ff:fe01:15-"int-PE-2-P-7"                       1/1/c3/1:1000
  -
    fe80::18:1ff:fe01:b-"int-PE-2-PE-3"(B)                    1/1/c1/1:1000
2001:db8:aaaa:102:0:2000::/128                SRV6            524289
  -
    fe80::24:1ff:fe01:1-"int-PE-2-P-6"                        1/1/c2/1:1000
  2001:db8:aaaa:106:0:1000::
    fe80::18:1ff:fe01:b-"int-PE-2-PE-3"(B)                    1/1/c1/1:1000
2001:db8:aaaa:102:0:3000::/128                SRV6            524290
  -
    fe80::28:1ff:fe01:15-"int-PE-2-P-7"                       1/1/c3/1:1000
  2001:db8:aaaa:107:0:1000::
    fe80::18:1ff:fe01:b-"int-PE-2-PE-3"(B)                    1/1/c1/1:1000
2001:db8:aaaa:102:0:4000::/128                SRV6            524291
  -
    fe80::18:1ff:fe01:b-"int-PE-2-PE-3"                       1/1/c1/1:1000
  2001:db8:aaaa:103:0:1000::
    fe80::24:1ff:fe01:1-"int-PE-2-P-6"(B)                     1/1/c2/1:1000
-------------------------------------------------------------------------------
Total Entries : 9
-------------------------------------------------------------------------------
===============================================================================
```

The IPv6 route table on PE-2 contains the following prefixes with shared block 2001:db8:aaaa::/48.

```
[/]
A:admin@PE-2# show router route-table ipv6 2001:db8:aaaa::/48 longer

===============================================================================
IPv6 Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
2001:db8:aaaa:101::/64                        Remote  ISIS      00h13m34s  18
       2001:db8:aaaa:101::/64 (tunneled:SRV6-ISIS)                 50
2001:db8:aaaa:102::/64                        Local   SRV6      00h22m44s  3
       fe80::201-"_tmnx_fpe_2.a"                                   0
2001:db8:aaaa:102:0:1000::/128                Local   SRV6      00h15m41s  3
```

```
       Black Hole                                                      0
2001:db8:aaaa:102:0:2000::/128               Local   ISIS      00h13m35s  18
       2001:db8:aaaa:102:0:2000:: (tunneled:SRV6-ISIS)                 10
2001:db8:aaaa:102:0:3000::/128               Local   ISIS      00h13m35s  18
       2001:db8:aaaa:102:0:3000:: (tunneled:SRV6-ISIS)                 10
2001:db8:aaaa:102:0:4000::/128               Local   ISIS      00h13m35s  18
       2001:db8:aaaa:102:0:4000:: (tunneled:SRV6-ISIS)                 10
2001:db8:aaaa:103::/64                       Remote  ISIS      00h13m27s  18
       2001:db8:aaaa:103::/64 (tunneled:SRV6-ISIS)                     20
2001:db8:aaaa:104::/64                       Remote  ISIS      00h13m21s  18
       2001:db8:aaaa:104::/64 (tunneled:SRV6-ISIS)                     40
2001:db8:aaaa:105::/64                       Remote  ISIS      00h13m15s  18
       2001:db8:aaaa:105::/64 (tunneled:SRV6-ISIS)                     30
2001:db8:aaaa:106::/64                       Remote  ISIS      00h13m10s  18
       2001:db8:aaaa:106::/64 (tunneled:SRV6-ISIS)                     20
2001:db8:aaaa:107::/64                       Remote  ISIS      00h13m02s  18
       2001:db8:aaaa:107::/64 (tunneled:SRV6-ISIS)                     20
-------------------------------------------------------------------------------
No. of Routes: 11
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
===============================================================================
```

The following command on PE-2 shows the corresponding FIB:

```
[/]
A:admin@PE-2# show router fib 1 ipv6 ip-prefix-prefix-length 2001:db8:aaaa::/48 longer

===============================================================================
FIB Display
===============================================================================
Prefix [Flags]                                          Protocol
  NextHop
-------------------------------------------------------------------------------
2001:db8:aaaa:101::/64                                  ISIS
  2001:db8:aaaa:101::/64 (Transport:SRV6:524292)
2001:db8:aaaa:102::/64                                  SRV6
  fe80::201 (_tmnx_fpe_2.a)
2001:db8:aaaa:102:0:1000::/128                          SRV6
  Blackhole
2001:db8:aaaa:102:0:2000::/128                          ISIS
  2001:db8:aaaa:102:0:2000:: (Transport:SRV6:524289)
2001:db8:aaaa:102:0:3000::/128                          ISIS
  2001:db8:aaaa:102:0:3000:: (Transport:SRV6:524290)
2001:db8:aaaa:102:0:4000::/128                          ISIS
  2001:db8:aaaa:102:0:4000:: (Transport:SRV6:524291)
2001:db8:aaaa:103::/64                                  ISIS
  2001:db8:aaaa:103::/64 (Transport:SRV6:524293)
2001:db8:aaaa:104::/64                                  ISIS
  2001:db8:aaaa:104::/64 (Transport:SRV6:524294)
2001:db8:aaaa:105::/64                                  ISIS
  2001:db8:aaaa:105::/64 (Transport:SRV6:524295)
2001:db8:aaaa:106::/64                                  ISIS
  2001:db8:aaaa:106::/64 (Transport:SRV6:524296)
2001:db8:aaaa:107::/64                                  ISIS
  2001:db8:aaaa:107::/64 (Transport:SRV6:524297)
-------------------------------------------------------------------------------
Total Entries : 11
-------------------------------------------------------------------------------
===============================================================================
```

Conclusion

SRv6 offers both shortest path and source routing capabilities. SRv6 can be deployed as an IPv6 transport for implementing services across a service provider network.

Segment Routing over IPv6 for VPRN

This chapter provides information about segment routing over IPv6 for VPRN.

Topics in this chapter include:

Applicability

The information and configuration in this chapter are based on SR OS Release 22.2.R1. Segment routing over IPv6 (SRv6) is supported on FP4-based equipment in SR OS Release 21.5.R2 and later.

Overview

SRv6 for VPRN allows the transport of VPRN-related IPv4 and IPv6 data across an SRv6-enabled network. To this end, VPRN-related data is sent to an ingress SRv6 router, where it is encapsulated and forwarded via an SRv6 tunnel. The SRv6 tunnel transports the encapsulated data across the SRv6-enabled network to an egress SRv6 router, where it is decapsulated and forwarded further as VPRN-related data. SRv6-tunneled data is encapsulated using an IPv6 header, where the destination address is a unique SRv6 segment identifier (SID), and is processed and forwarded in the IPv6 data plane.

An SRv6 SID is a preconfigured 128-bit routable IPv6 prefix address that is encoded in three parts: a locator, a function, and an argument. The locator is a summary IPv6 prefix for a set of SRv6 SIDs instantiated on an SRv6-capable router. It is used to route the data within the IPv6 transport network. Each participating SRv6-capable router needs its unique locator, based on a common block that all participating SRv6-capable routers share in the IPv6 address space. The function is an opaque identifier that indicates the local behavior at the endpoint of an SRv6 segment. The focus in this topic is on the SRv6 End.DT4 and the SRv6 End.DT6 functions for the VPRN, performing a prefix lookup in the VPRN service IPv4 route table (End.DT4) or in the VPRN service IPv6 route table (End.DT6). The argument is not used in SR OS 22.2.R1 and is set to all zeros.

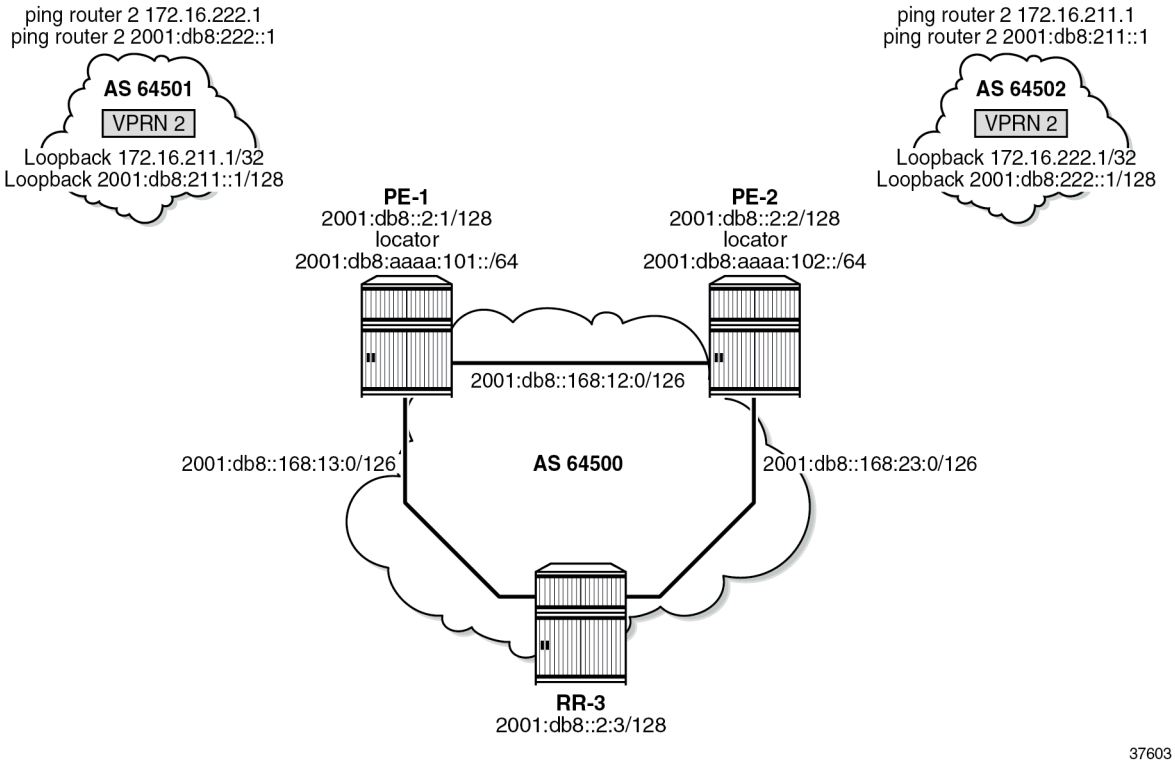
The local router installs its locator prefix in its IPv6 route table and forwarding information base (FIB), and advertises its locator prefix in IS-IS with the SRv6 locator sub-TLV. Each remote router populates its route table and FIB with the received locator prefixes, including the tunneled next hop to the originating router. Each remote router also populates its VPRN service route table with the received network prefixes, including the tunneled next hop to the VPRN of the originating router.

SRv6 data transport requires additional processing at both the ingress and egress data planes. This processing relies on forwarding path extension (FPE), as described in the Segment Routing over IPv6 chapter.

Configuration

Figure 55: Example topology shows the example topology with three routers. The SRv6-enabled network that it represents comprises PE-1, PE-2, and a route reflector RR-3 in the control plane. The SRv6-enabled network has only IPv6 addresses and interfaces.

Figure 55: Example topology



For the transport of IPv4 and IPv6 data from the VPRN on PE-1 to the VPRN on PE-2, PE-1 acts as the SRv6 ingress PE node, while PE-2 acts as the SRv6 egress PE node. For the transport of IPv4 and IPv6 data from the VPRN on PE-2 to the VPRN on PE-1, PE-2 acts as the SRv6 ingress PE node, while PE-1 acts as the SRv6 egress PE node. To explain SRv6 for VPRN, the topology does not need an SRv6 transit router, because SRv6 transit routers simply forward SRv6-encapsulated packets via IPv6 route table lookup without any other processing.

SRv6 and FPE are configured only on PE-1 and on PE-2. RR-3 acts as the BGP route reflector in the control plane. RR-3 does not participate in the SRv6 data transport that only exists between PE-1 and PE-2.

The **ping** and **traceroute** commands between IPv4 and IPv6 loopback addresses in the VPRNs simulate data transport.

The configuration for this example topology is symmetrical. All **configure** and **show** command output examples for PE-1 also apply to PE-2. The **configure** and **show** commands with deviating output examples for RR-3 are explicitly mentioned.

Configure the router

This configuration includes:

- ports and IPv6-only interfaces on PE-1, PE-2, and RR-3
- port cross-connect (PXC) on PE-1 and PE-2, using internal loopbacks on an FP4 MAC chip, as described in the Segment Routing over IPv6 chapter
- IS-IS
  - On PE-1, PE-2, and RR-3, include:
    - level 2 capability with wide metrics (for the 128-bit identifiers)
    - native IPv6 routing
  - On PE-1 and PE-2, as a best practice to advertise the router capability within the autonomous system (AS), also configure:
    - **traffic-engineering**
    - **traffic-engineering-options**
- BGP on PE-1, PE-2, and RR-3, with internal group “gr_v6_internal” that includes:
  - IPv4 and IPv6 families
  - **extended-nh-encoding** for IPv4
  - **advertise-ipv6-next-hops** for IPv4
  - BGP neighbor **system** IPv6 addresses
  - On PE-1 and PE-2 only: **next-hop-self**

The core network topology uses IPv6 for BGP peering (with 16 byte next hop addresses), so to advertise and receive IPv4 routes (which have 4 byte next hop addresses) with IPv6 next hop addresses, the commands **advertise-ipv6-next-hops** and **extended-nh-encoding** need to be configured at the BGP, group, or neighbor level. The **advertise-ipv6-next-hops** command instructs the system to advertise IPv4 routes with IPv6 next hop addresses. The **extended-nh-encoding** command configures BGP to advertise the capability to receive IPv4 routes with IPv6 next hop addresses.

The following example configuration applies for PE-1 and is similar for PE-2.

```
[/]
A:admin@PE-1# configure {
    router "Base" {
        autonomous-system 64500
        interface "int-PE-1-PE-2" {
            description "interface between PE-1 and PE-2"
            port 1/1/c1/1:1000
            ipv6 {
                address 2001:db8::168:12:1 {
                    prefix-length 126
                }
            }
        }
        interface "int-PE-1-RR-3" {
            description "interface between PE-1 and RR-3"
            port 1/1/c2/1:1000
            ipv6 {
                address 2001:db8::168:13:1 {
                    prefix-length 126
```

```
            }
        }
    }
    interface "system" {
        description "system interface of PE-1"
        ipv6 {
            address 2001:db8::2:1 {
                prefix-length 128
            }
        }
    }
    bgp {
        min-route-advertisement 1
        router-id 2.2.2.1
        rapid-withdrawal true
        split-horizon true
        ebgp-default-reject-policy {
            import false
            export false
        }
        group "gr_v6_internal" {
            description "internal bgp group on PE-1"
            next-hop-self true
            type internal
            family {
                ipv4 true
                ipv6 true
            }
            extended-nh-encoding {
                ipv4 true
            }
            advertise-ipv6-next-hops {
                ipv4 true
            }
        }
        neighbor "2001:db8::2:3" {    # RR-3 system address
            group "gr_v6_internal"
        }
    }
    isis 0 {
        admin-state enable
        advertise-router-capability as
        ipv6-routing native
        level-capability 2    # required for SRv6
        router-id 1.1.1.1
        traffic-engineering true
        area-address [49.0001]
        traffic-engineering-options {
            ipv6 true
            application-link-attributes {
            }
        }
        interface "int-PE-1-PE-2" {
            interface-type point-to-point
        }
        interface "int-PE-1-RR-3" {
            interface-type point-to-point
        }
        interface "system" {
            passive true
        }
        level 2 {
            wide-metrics-only true    # required for SRv6
        }
```

```
        }
        exit all
```

The following example configuration applies for RR-3:

```
[/]
A:admin@RR-3# configure {
    router "Base" {
        autonomous-system 64500
        interface "int-RR-3-PE-1" {
            description "interface between RR-3 and PE-1"
            port 1/1/c1/1:1000
            ipv6 {
                address 2001:db8::168:13:2 {
                    prefix-length 126
                }
            }
        }
        interface "int-RR-3-PE-2" {
            description "interface between RR-3 and PE-2"
            port 1/1/c2/1:1000
            ipv6 {
                address 2001:db8::168:23:2 {
                    prefix-length 126
                }
            }
        }
        interface "system" {
            description "system interface of RR-3"
            ipv6 {
                address 2001:db8::2:3 {
                    prefix-length 128
                }
            }
        }
        bgp {
            min-route-advertisement 1
            router-id 2.2.2.3
            rapid-withdrawal true
            split-horizon true
            ebgp-default-reject-policy {
                import false
                export false
            }
            group "gr_v6_internal" {
                description "internal bgp group on RR-3"
                type internal
                family {
                    ipv4 true
                    ipv6 true
                }
                cluster {
                    cluster-id 3.3.3.3
                }
                extended-nh-encoding {
                    ipv4 true
                }
                advertise-ipv6-next-hops {
                    ipv4 true
                }
            }
            neighbor "2001:db8::2:1" {    # PE-1 system address
                group "gr_v6_internal"
```

```
            }
            neighbor "2001:db8::2:2" {    # PE-2 system address
                group "gr_v6_internal"
            }
        }
        isis 0 {
            admin-state enable
            ipv6-routing native
            level-capability 2    # required for SRv6
            router-id 1.1.1.3
            area-address [49.0001]
            interface "int-RR-3-PE-1" {
                interface-type point-to-point
            }
            interface "int-RR-3-PE-2" {
                interface-type point-to-point
            }
            interface "system" {
                passive true
            }
            level 2 {
                wide-metrics-only true    # required for SRv6
            }
        }
        exit all
```

Configure the VPRN services on PE-1 and on PE-2

This configuration includes:

- an IPv4 address and an IPv6 address for a loopback interface "lb_itf_vprn"
- BGP, with external group “gr_v6_vprn” that includes the following capabilities:
  - IPv4 and IPv6 families
  - **extended-nh-encoding** for IPv4
  - **advertise-ipv6-next-hops** for IPv4
  - BGP neighbor **interface** IPv6 addresses, with BGP neighbors in a different external AS

The following example configuration applies for VPRN 2 on PE-1 and is similar for VPRN 2 on PE-2.

```
[/]
A:admin@PE-1# configure {
    service {
        vprn "VPRN_2" {
            admin-state enable
            description "VPRN 2 on PE-1"
            service-id 2
            customer "1"
            autonomous-system 64500
            bgp {
                ebgp-default-reject-policy {
                    import false
                    export false
                }
                group "gr_v6_vprn" {
                    description "external bgp group for VPRN 2 on PE-1"
                    family {
                        ipv4 true
```

```
                    ipv6 true
                }
                extended-nh-encoding {
                    ipv4 true
                }
                advertise-ipv6-next-hops {
                    ipv4 true
                }
            }
            neighbor "2001:db8:101::1" {
                group "gr_v6_vprn"
                type external
                peer-as 64501
            }
        }
        interface "lb_itf_vprn" {
            description "VPRN 2 interface on PE-1 for external subnet"
            loopback true
            ipv4 {
                primary {
                    address 172.16.211.1
                    prefix-length 32
                }
            }
            ipv6 {
                address 2001:db8:211::1 {
                    prefix-length 128
                }
            }
        }
        exit all
```

At this point, verify that data transport is not possible between the local VPRN on PE-1 and the remote VPRN on PE-2.

```
[/]
A:admin@PE-1# ping 172.16.222.1 router-instance "VPRN_2"
PING 172.16.222.1 56 data bytes
No route to destination. Address: 172.16.222.1, Router Instance: "VPRN_2"
---snip---
---- 172.16.222.1 PING Statistics ----
5 packets transmitted, 0 packets received, 100% packet loss

[/]
A:admin@PE-1# ping 2001:db8:222::1 router-instance "VPRN_2"
PING 2001:db8:222::1 56 data bytes
No route to destination. Address: 2001:db8:222::1, Router Instance: "VPRN_2"
---snip---
---- 2001:db8:222::1 PING Statistics ----
5 packets transmitted, 0 packets received, 100% packet loss
```

The result of the verification complies with the route table for the local VPRN on PE-1 that only contains local routes for its own loopback addresses:

```
[/]
A:admin@PE-1# show router 2 route-table ipv4

===============================================================================
Route Table (Service: 2)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                         Active     Metric
```

```
-------------------------------------------------------------------------------
172.16.211.1/32                                 Local   Local     00h01m21s  0
       lb_itf_vprn                                       Y           0
-------------------------------------------------------------------------------
No. of Routes: 1
---snip---
===============================================================================
```

```
[/]
A:admin@PE-1# show router 2 route-table ipv6

===============================================================================
IPv6 Route Table (Service: 2)
===============================================================================
Dest Prefix[Flags]                              Type    Proto     Age        Pref
      Next Hop[Interface Name]                            Active     Metric
-------------------------------------------------------------------------------
2001:db8:211::1/128                             Local   Local     00h01m19s  0
       lb_itf_vprn                                       Y           0
-------------------------------------------------------------------------------
No. of Routes: 1
---snip---
===============================================================================
```

Perform the same verification for data transport between the remote VPRN on PE-2 and the local VPRN on PE-1.

Configure SRv6 in the router Base context on PE-1 and on PE-2

Configure the locator in the **router Base segment-routing segment-routing-v6** context on PE-2 and similar on PE-1, with **ip-prefix** 2001:db8:aaaa:101::/64 for locator “PE-1_loc”.

```
[/]
A:admin@PE-2# configure {
    router "Base" {
        segment-routing {
            segment-routing-v6 {
                locator "PE-2_loc" {
                    admin-state enable
                    block-length 48
                    prefix {
                        ip-prefix 2001:db8:aaaa:102::/64
                    }
                }
                exit all
```

Configure the FPEs on PE-2 and identical on PE-1.

```
[/]
A:admin@PE-2# configure {
    fwd-path-ext {
        fpe 1 {
            path {
                pxc 1
            }
            application {
                srv6 {
                    type origination
                }
```

```
            }
        }
        fpe 2 {
            path {
                pxc 2
            }
            application {
                srv6 {
                    type termination
                }
            }
        }
        exit all
```

Use FPE 1 as the SRv6 origination FPE in the **router Base segment-routing segment-routing-v6** context and FPE 2 as the SRv6 termination FPE in the **router Base segment-routing segment-routing-v6 locator** context on PE-2 and similar on PE-1, for locator “PE-1_loc”. For more information, see the Segment Routing over IPv6 chapter.

```
[/]
A:admin@PE-2# configure {
    router "Base" {
        segment-routing {
            segment-routing-v6 {
                origination-fpe [1]
                locator "PE-2_loc" {
                    admin-state enable
                    termination-fpe [2]
                }
                exit all
```

Configure the SRv6 End function (equivalent to an IPv4 node SID) and SRv6 End.X functions (equivalent to IPv4 adjacency SIDs) in the **router Base segment-routing segment-routing-v6 base-routing-instance locator** context on PE-2 and similar on PE-1, for locator “PE-1_loc”.

```
[/]
A:admin@PE-2# configure {
    router "Base" {
        segment-routing {
            segment-routing-v6 {
                base-routing-instance {
                    locator "PE-2_loc" {
                        function {
                            end 1 {
                                srh-mode usp
                            }
                            end-x-automatically allocate psp protection unprotected { }
                        }
                    }
                    exit all
```

Advertise the locator in IS-IS while ensuring level 2 capability on PE-2 and similar on PE-1, for locator “PE-1_loc”.

```
[/]
A:admin@PE-2# configure {
    router "Base" {
        isis 0 {
            segment-routing-v6 {
                admin-state enable
```

```
                locator "PE-2_loc" {
                    level-capability 2
                }
            }
            exit all
```

A summary of the locator and origination FPE configuration can be displayed with the **show router segment-routing-v6 summary** command.

Verify the SRv6 local SIDs on PE-2 and similar on PE-1. Three SRv6 local SIDs are created: one for the statically configured SRv6 End function (configured in the **router Base segment-routing segment-routing-v6 base-routing-instance locator** context) and two for the automatically allocated SRv6 End.X functions (one facing PE-1 and one facing RR-3). All three SRv6 local SIDs are concatenated with the locator. The statically configured SRv6 End function appears first with function number 1. In the example, the automatically allocated SRv6 End.X functions receive function numbers 2 and 4 respectively. RR-3 has no SRv6 configuration and does not have these SRv6 local SIDs and SRv6 functions.

```
[/]
A:admin@PE-2# show router segment-routing-v6 local-sid

===============================================================================
Segment Routing v6 Local SIDs
===============================================================================
SID                                                 Type           Function
  Locator
  Context
-------------------------------------------------------------------------------
2001:db8:aaaa:102:0:1000::                          End            1
  PE-2_loc
  Base
2001:db8:aaaa:102:0:2000::                          End.X          2
  PE-2_loc
    None
2001:db8:aaaa:102:0:4000::                          End.X          4
  PE-2_loc
    None
-------------------------------------------------------------------------------
SIDs : 3
-------------------------------------------------------------------------------
===============================================================================
```

Verify the SRv6 base routing instance details on PE-2 and similar on PE-1. The SRv6 functions for the configured locator are listed. The SRv6 End function is statically configured. There is an automatically allocated SRv6 End.X function for each IS-IS neighbor.

```
[/]
A:admin@PE-2# show router segment-routing-v6 base-routing-instance

===============================================================================
Segment Routing v6 Base Routing Instance
===============================================================================
Locator
  Type       Function    SID                                     Status/InstId
    SRH-mode Protection  Interface
-------------------------------------------------------------------------------
PE-2_loc
  End                 1 2001:db8:aaaa:102:0:1000::              ok
    USP
-------------------------------------------------------------------------------
  Auto-allocated End.X: PSP Unprotected,
-------------------------------------------------------------------------------
```

```
  End.X              *2 2001:db8:aaaa:102:0:2000::             0
    PSP      Unprotected int-PE-2-PE-1
    ISIS Level: L2 Mac Address: 04:0a:01:01:00:01 Nbr Sys Id: 0010.0100.1001
  End.X              *4 2001:db8:aaaa:102:0:4000::             0
    PSP      Unprotected int-PE-2-RR-3
    ISIS Level: L2 Mac Address: 04:12:01:01:00:0b Nbr Sys Id: 0010.0100.1003
-------------------------------------------------------------------------------
===============================================================================
Legend: * - System allocated
```

Verify the IPv6 route table on PE-1. The IPv6 route table also has routes to the local and learned remote locators and to the local SRv6 function SIDs. The remotely configured locator prefix of PE-2 is reached via an SRv6 tunnel. The routes with protocol “SRV6” correspond with the locally configured locator prefix of PE-1 or the locally configured SRv6 End function.

```
[/]
A:admin@PE-1# show router route-table ipv6

===============================================================================
IPv6 Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
2001:db8::2:1/128                             Local   Local     00h10m45s  0
       system                                                        0
2001:db8::2:2/128                             Remote  ISIS      00h00m27s  18
       fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                          10
2001:db8::2:3/128                             Remote  ISIS      00h00m27s  18
       fe80::612:1ff:fe01:1-"int-PE-1-RR-3"                          10
2001:db8::168:12:0/126                        Local   Local     00h10m44s  0
       int-PE-1-PE-2                                                 0
2001:db8::168:13:0/126                        Local   Local     00h10m44s  0
       int-PE-1-RR-3                                                 0
2001:db8::168:23:0/126                        Remote  ISIS      00h00m27s  18
       fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                          20
2001:db8:aaaa:101::/64                        Local   SRV6      00h01m47s  3
       fe80::201-"_tmnx_fpe_2.a"                                     0
2001:db8:aaaa:101:0:1000::/128                Local   SRV6      00h01m01s  3
       Black Hole                                                    0
2001:db8:aaaa:101:0:2000::/128                Local   ISIS      00h00m28s  18
       fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                          10
2001:db8:aaaa:101:0:4000::/128                Local   ISIS      00h00m28s  18
       fe80::612:1ff:fe01:1-"int-PE-1-RR-3"                          10
2001:db8:aaaa:102::/64                        Remote  ISIS      00h00m17s  18
       2001:db8:aaaa:102::/64 (tunneled:SRV6-ISIS)                   10
-------------------------------------------------------------------------------
No. of Routes: 11
---snip---
===============================================================================
```

Verify that the tunnel from PE-1 to the remote locator has SRv6 encapsulation and similar for the tunnel from PE-2. The tunnel table on RR-3 remains empty.

```
[/]
A:admin@PE-1# show router tunnel-table ipv6

===============================================================================
IPv6 Tunnel Table (Router: Base)
===============================================================================
Destination                                   Owner     Encap TunnelId  Pref
Nexthop                                       Color           Metric
```

```
-------------------------------------------------------------------------------
2001:db8:aaaa:102::/64                          srv6-isis SRV6  524289    0
  fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                            10
-------------------------------------------------------------------------------
---snip---
===============================================================================
```

Verify that the tunnel from PE-1 to the remote locator uses the “int-PE-1-PE-2” interface and similar for the tunnel from PE-2, where the tunnel to the remote locator uses the “int-PE-2-PE-1” interface. Interface “int-PE-1-PE-2” is configured on port 1/1/c1/1:1000. The FP tunnel table on RR-3 remains empty.

```
[/]
A:admin@PE-1# show router fp-tunnel-table 1 ipv6

===============================================================================
IPv6 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                Protocol         Tunnel-ID
  Lbl/SID
    NextHop                                               Intf/Tunnel
  Lbl/SID (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
2001:db8:aaaa:102::/64                     SRV6              524289
  -
    fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                    1/1/c1/1:1000
-------------------------------------------------------------------------------
Total Entries : 1
-------------------------------------------------------------------------------
===============================================================================
```

Verify the IS-IS data base on PE-1 with **show router isis database detail**. The output of this command provides information on each IS-IS-enabled router. Per uniquely identified IS-IS-enabled router, the SRv6 information indicates:

- the IS-IS-advertised router capabilities
- the advertised SRv6 locator TLV
- the advertised configured SRv6 End SID and automatically allocated SRv6 End.X SIDs

```
[/]
A:admin@PE-1# show router isis database detail

===============================================================================
Rtr Base ISIS Instance 0 Database (detail)
===============================================================================

Displaying Level 1 database
-------------------------------------------------------------------------------
Level (1) LSP Count : 0

Displaying Level 2 database
-------------------------------------------------------------------------------
LSP ID    : PE-1.00-00                                  Level     : L2
Sequence  : 0x6                  Checksum  : 0x97fb   Lifetime  : 1189
Version   : 1                    Pkt Type  : 20       Pkt Ver   : 1
Attributes: L1L2                 Max Area  : 3        Alloc Len : 1492
```

```
SYS ID    : 0010.0100.1001         SysID Len : 6        Used Len  : 398

TLVs :
  Area Addresses:
    Area Address : (3) 49.0001
  Supp Protocols:
    Protocols     : IPv4
    Protocols     : IPv6
  IS-Hostname   : PE-1
  Router ID   :
    Router ID   : 1.1.1.1
  TE Router ID v6  :
    Router ID   : 2001:db8::2:1
  Router Cap : 1.1.1.1, D:0, S:0
    TE Node Cap : B E M  P
    SRv6 Cap: 0x0000
    SR Alg: metric based SPF
    Node MSD Cap: BMI : 0 SRH-MAX-SL : 10 SRH-MAX-END-POP : 9 SRH-MAX-H-ENCAPS : 1 SRH-MAX-END-
D : 9
  I/F Addresses IPv6 :
    IPv6 Address    : 2001:db8::2:1
    IPv6 Address    : 2001:db8::168:12:1
    IPv6 Address    : 2001:db8::168:13:1
  TE IS Nbrs   :
    Nbr   : PE-2.00
    Default Metric  : 10
    Sub TLV Len     : 60
    IPv6 Addr : 2001:db8::168:12:1
    Nbr IPv6  : 2001:db8::168:12:2
    End.X-SID: 2001:db8:aaaa:101:0:2000:: flags: algo:0 weight:0 endpoint:End.X-PSP
  TE IS Nbrs   :
    Nbr   : RR-3.00
    Default Metric  : 10
    Sub TLV Len     : 42
    IPv6 Addr : 2001:db8::168:13:1
    End.X-SID: 2001:db8:aaaa:101:0:4000:: flags: algo:0 weight:0 endpoint:End.X-PSP
  IPv6 Reach:
    Metric: ( I ) 0
    Prefix   : 2001:db8::2:1/128
    Metric: ( I ) 10
    Prefix   : 2001:db8::168:12:0/126
    Metric: ( I ) 10
    Prefix   : 2001:db8::168:13:0/126
    Metric: ( I ) 0
    Prefix   : 2001:db8:aaaa:101::/64
  SRv6 Locator  :
    MT ID : 0
    Metric: ( ) 0 Algo:0
    Prefix   : 2001:db8:aaaa:101::/64
    Sub TLV   :
      End-SID   : 2001:db8:aaaa:101:0:1000::, flags:0x0, endpoint:End-USP

-------------------------------------------------------------------------------
LSP ID    : PE-2.00-00                                 Level     : L2
Sequence  : 0x6                  Checksum  : 0x1740   Lifetime  : 1185
Version   : 1                    Pkt Type  : 20       Pkt Ver   : 1
Attributes: L1L2                 Max Area  : 3        Alloc Len : 398
SYS ID    : 0010.0100.1002       SysID Len : 6        Used Len  : 398

TLVs :
  Area Addresses:
    Area Address : (3) 49.0001
  Supp Protocols:
    Protocols     : IPv4
```

```
    Protocols     : IPv6
  IS-Hostname   : PE-2
  Router ID   :
    Router ID   : 1.1.1.2
  TE Router ID v6  :
    Router ID   : 2001:db8::2:2
  Router Cap : 1.1.1.2, D:0, S:0
    TE Node Cap : B E M  P
    SRv6 Cap: 0x0000
    SR Alg: metric based SPF
    Node MSD Cap: BMI : 0 SRH-MAX-SL : 10 SRH-MAX-END-POP : 9 SRH-MAX-H-ENCAPS : 1 SRH-MAX-END-
D : 9
  I/F Addresses IPv6 :
    IPv6 Address    : 2001:db8::2:2
    IPv6 Address    : 2001:db8::168:12:2
    IPv6 Address    : 2001:db8::168:23:1
  TE IS Nbrs   :
    Nbr   : PE-1.00
    Default Metric  : 10
    Sub TLV Len     : 60
    IPv6 Addr : 2001:db8::168:12:2
    Nbr IPv6  : 2001:db8::168:12:1
    End.X-SID: 2001:db8:aaaa:102:0:2000:: flags: algo:0 weight:0 endpoint:End.X-PSP
  TE IS Nbrs   :
    Nbr   : RR-3.00
    Default Metric  : 10
    Sub TLV Len     : 42
    IPv6 Addr : 2001:db8::168:23:1
    End.X-SID: 2001:db8:aaaa:102:0:4000:: flags: algo:0 weight:0 endpoint:End.X-PSP
  IPv6 Reach:
    Metric: ( I ) 0
    Prefix   : 2001:db8::2:2/128
    Metric: ( I ) 10
    Prefix   : 2001:db8::168:12:0/126
    Metric: ( I ) 10
    Prefix   : 2001:db8::168:23:0/126
    Metric: ( I ) 0
    Prefix   : 2001:db8:aaaa:102::/64
  SRv6 Locator  :
    MT ID : 0
    Metric: ( ) 0 Algo:0
    Prefix   : 2001:db8:aaaa:102::/64
    Sub TLV   :
      End-SID   : 2001:db8:aaaa:102:0:1000::, flags:0x0, endpoint:End-USP

-------------------------------------------------------------------------------
LSP ID    : RR-3.00-00                                Level     : L2
Sequence  : 0x3                  Checksum  : 0xdba6   Lifetime  : 675
Version   : 1                    Pkt Type  : 20       Pkt Ver   : 1
Attributes: L1L2                 Max Area  : 3        Alloc Len : 193
SYS ID    : 0010.0100.1003       SysID Len : 6        Used Len  : 193

TLVs :
  Area Addresses:
    Area Address : (3) 49.0001
  Supp Protocols:
    Protocols     : IPv4
    Protocols     : IPv6
  IS-Hostname   : RR-3
  Router ID   :
    Router ID   : 1.1.1.3
  I/F Addresses IPv6 :
    IPv6 Address    : 2001:db8::2:3
    IPv6 Address    : 2001:db8::168:13:2
```

```
    IPv6 Address    : 2001:db8::168:23:2
  TE IS Nbrs   :
    Nbr   : PE-1.00
    Default Metric  : 10
    Sub TLV Len     : 0
  TE IS Nbrs   :
    Nbr   : PE-2.00
    Default Metric  : 10
    Sub TLV Len     : 0
  IPv6 Reach:
    Metric: ( I ) 0
    Prefix   : 2001:db8::2:3/128
    Metric: ( I ) 10
    Prefix   : 2001:db8::168:13:0/126
    Metric: ( I ) 10
    Prefix   : 2001:db8::168:23:0/126

Level (2) LSP Count : 3
-------------------------------------------------------------------------------
---snip---
===============================================================================
```

Verify the IS-IS routes on PE-1 and similar on PE-2.

```
[/]
A:admin@PE-1# show router isis routes

===============================================================================
Rtr Base ISIS Instance 0 Route Table
===============================================================================
Prefix[Flags]                     Metric     Lvl/Typ     Ver.  SysID/Hostname
  NextHop                                                MT     AdminTag/SID[F]
-------------------------------------------------------------------------------
2001:db8::2:1/128                 0          2/Int.      2     PE-1
   ::                                                      0       0
2001:db8::2:2/128                 10         2/Int.      10    PE-2
   fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                    0       0
2001:db8::2:3/128                 10         2/Int.      10    RR-3
   fe80::612:1ff:fe01:1-"int-PE-1-RR-3"                    0       0
2001:db8::168:12:0/126            10         2/Int.      4     PE-1
   ::                                                      0       0
2001:db8::168:13:0/126            10         2/Int.      4     PE-1
   ::                                                      0       0
2001:db8::168:23:0/126            20         2/Int.      10    PE-2
   fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                    0       0
2001:db8:aaaa:101::/64            0          2/Int.      12    PE-1
   ::                                                      0       0
2001:db8:aaaa:102::/64            10         2/Int.      11    PE-2
   fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                    0       0
-------------------------------------------------------------------------------
No. of Routes: 8 (8 paths)
-------------------------------------------------------------------------------
---snip---
===============================================================================
```

This output corresponds with the information in the route table and in the FIB.

The BGP groups can be verified with the **show router bgp group** command. PE-1 and PE-2 know the internal and the external BGP groups. RR-3 only knows the internal BGP group.

The BGP next hops can be verified with the following commands:

- **show router bgp next-hop ipv4**

- **show router bgp next-hop ipv6**
- **show router bgp next-hop vpn-ipv4**
- **show router bgp next-hop vpn-ipv6**

Verify on PE-1 and similar on PE-2 that the locator prefixes are locally configured and advertised. In this example, PE-1 is aware of both locators. One locator is locally configured; the other is learned from the PE-2 advertisement.

```
[/]
A:admin@PE-1# show router isis segment-routing-v6 locator

===============================================================================
Rtr Base ISIS Instance 0 SRv6 Locator Table
===============================================================================
Prefix                            AdvRtr                        MT     Lvl/Typ
 AttributeFlags                    Tag                           Flags  Algo
-------------------------------------------------------------------------------
2001:db8:aaaa:101::/64            PE-1                          0      2/Int.
  -                                 0                             -      0
2001:db8:aaaa:102::/64            PE-2                          0      2/Int.
  -                                 0                             -      0
-------------------------------------------------------------------------------
No. of Locators: 2
-------------------------------------------------------------------------------
---snip---
===============================================================================
```

Verify on PE-1 and similar on PE-2 that the SRv6 End SIDs are locally configured and advertised. In this example, PE-1 is aware of both SRv6 End SIDs. One End SID is locally configured; the other is learned from the PE-2 advertisement.

```
A:admin@PE-1# show router isis segment-routing-v6 end-sid

===============================================================================
Rtr Base ISIS Instance 0 SRv6 End SID Table
===============================================================================
Prefix                            AdvRtr                        MT     Lvl/Typ
 Sid                               Behavior                      Flags  Algo
-------------------------------------------------------------------------------
2001:db8:aaaa:101::/64            PE-1                          0      2/Int.
  2001:db8:aaaa:101:0:1000::        End USP                       -      0
2001:db8:aaaa:102::/64            PE-2                          0      2/Int.
  2001:db8:aaaa:102:0:1000::        End USP                       -      0
-------------------------------------------------------------------------------
No. of End SIDs: 2
===============================================================================
```

Configure SRv6 for the VPRNs on PE-1 and on PE-2

On PE-1, PE-2, and RR-3, extend the BGP advertisements to include the VPN-IPv4 and VPN-IPv6 families.

```
configure {
    router "Base" {
        bgp {
            rapid-update {
                vpn-ipv4 true
```

```
                vpn-ipv6 true
            }
            group "gr_v6_internal" {
                family {
                    ipv4 true
                    vpn-ipv4 true
                    ipv6 true
                    vpn-ipv6 true
                }
                extended-nh-encoding {
                    vpn-ipv4 true
                    ipv4 true
                }
                advertise-ipv6-next-hops {
                    vpn-ipv6 true
                    vpn-ipv4 true
                    ipv4 true
                }
            }
            exit all
```

On PE-2, create an SRv6 instance for the VPRN service. Use the locator from the **router Base segment-routing segment-routing-v6** context and configure End.DT4 and End.DT6 functions for it.

Use the created SRv6 instance in the **service vprn bgp-ipvpn segment-routing-v6** context, with the configured locator as the default locator. Ensure a unique route distinguisher. Use the unique PE-2 system IPv6 address as the source address. Perform a similar configuration on PE-1, with the PE-1 locator as the default locator, the PE-1 system IPv6 address as the source address, and a different route distinguisher.

```
[/]
A:admin@PE-2# configure {
    service {
        vprn "VPRN_2" {
            segment-routing-v6 1 {
                locator "PE-2_loc" {
                    function {
                        end-dt4 {
                        }
                        end-dt6 {
                        }
                    }
                }
            }
            bgp-ipvpn {
                segment-routing-v6 1 {
                    admin-state enable
                    route-distinguisher "192.0.2.2:2"
                    source-address 2001:db8::2:2
                    vrf-target {
                        community "target:64506:2"
                    }
                    srv6 {
                        instance 1
                        default-locator "PE-2_loc"
                    }
                }
            }
            exit all
```

This configuration results in BGP update exchanges from PE-2 to PE-1, via RR-3, and similar from PE-1 to PE-2, via RR-3. PE-2 sends BGP updates to RR-3 for the VPN-IPv4 and the VPN-IPv6 families respectively. Each BGP update advertises the VPN-IPv4 or VPN-IPv6 address family, the reachable

network prefixes, the AS to which they belong, and an SRv6 Services TLV. The SRv6 Services TLV indicates that resolution to an SRv6 SID is available, making use of the endpoint behavior that is configured for the VPN-IPv4 or VPN-IPv6 address family on the locator. PE-1 programs the route prefixes with an SRv6 tunnel next hop in its VPRN service route table and in its FIB. PE-1 and PE-2 advertise only the SRv6 SIDs for the SRv6 End.DT4 and SRv6 End.DT6 functions.

When debug logging for BGP updates is configured, this configuration results in the following BGP update logs for the VPN-IPv4 address family.

Consider the example for VPN-IPv4 prefix 172.16.222.1/32. Similar BGP update logs are generated also for VPN-IPv4 prefix 172.16.211.1/32, in the other direction.

The following BGP update log is for the VPN-IPv4 address family. It is sent by PE-2 and received (via RR-3) by PE-1:

```
[/]
A:admin@PE-1# show log log-id "log_2"

---snip---
3 2022/06/21 15:07:32.057 CEST MINOR: DEBUG #2001 Base Peer 1: 2001:db8::2:3
"Peer 1: 2001:db8::2:3: UPDATE
Peer 1: 2001:db8::2:3 - Received BGP UPDATE:
    Withdrawn Length = 0
    Total Path Attr Length = 128
    Flag: 0x90 Type: 14 Len: 45 Multiprotocol Reachable NLRI:
        Address Family VPN_IPV4
        NextHop len 24 NextHop 2001:db8::2:2
        172.16.222.1/32 RD 192.0.2.2:2 Label 524288 (Raw label 0x800001)
    Flag: 0x40 Type: 1 Len: 1 Origin: 0
    Flag: 0x40 Type: 2 Len: 0 AS Path:
    Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
    Flag: 0x80 Type: 9 Len: 4 Originator ID: 2.2.2.2
    Flag: 0x80 Type: 10 Len: 4 Cluster ID:
        3.3.3.3
    Flag: 0xc0 Type: 16 Len: 8 Extended Community:
        target:64506:2
    Flag: 0xc0 Type: 40 Len: 37 Prefix-SID-attr:
       SRv6 Services TLV (37 bytes):-
           Type: SRV6 L3 Service TLV (5)
           Length: 34 bytes, Reserved: 0x0
         SRv6 Service Information Sub-TLV (33 bytes)
             Type: 1 Len: 30 Rsvd1: 0x0
             SRv6 SID: 2001:db8:aaaa:102::
             SID Flags: 0x0 Endpoint Behavior: 0x13 Rsvd2: 0x0
             SRv6 SID Sub-Sub-TLV
                Type: 1 Len: 6
                BL:48 NL:16 FL:20 AL:0 TL:20 TO:64
"
---snip---
```

Similar BGP update logs are generated for the VPN-IPv6 address family.

Consider the example for VPN-IPv6 prefix 2001:db8:222::1/128. Similar BGP update logs are generated also for VPN-IPv6 prefix 2001:db8:211::1/128, in the other direction.

The following BGP update log is for the VPN-IPv6 address family. It is sent by PE-2 and received (via RR-3) by PE-1:

```
[/]
A:admin@PE-1# show log log-id "log_2"

---snip---
```

```
4 2022/06/21 15:07:32.057 CEST MINOR: DEBUG #2001 Base Peer 1: 2001:db8::2:3
"Peer 1: 2001:db8::2:3: UPDATE
Peer 1: 2001:db8::2:3 - Received BGP UPDATE:
    Withdrawn Length = 0
    Total Path Attr Length = 140
    Flag: 0x90 Type: 14 Len: 57 Multiprotocol Reachable NLRI:
        Address Family VPN_IPV6
        NextHop len 24 NextHop 2001:db8::2:2
        2001:db8:222::1/128 RD 192.0.2.2:2 Label 524287 (Raw label 0x7ffff1)
    Flag: 0x40 Type: 1 Len: 1 Origin: 0
    Flag: 0x40 Type: 2 Len: 0 AS Path:
    Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
    Flag: 0x80 Type: 9 Len: 4 Originator ID: 2.2.2.2
    Flag: 0x80 Type: 10 Len: 4 Cluster ID:
        3.3.3.3
    Flag: 0xc0 Type: 16 Len: 8 Extended Community:
        target:64506:2
    Flag: 0xc0 Type: 40 Len: 37 Prefix-SID-attr:
       SRv6 Services TLV (37 bytes):-
           Type: SRV6 L3 Service TLV (5)
           Length: 34 bytes, Reserved: 0x0
         SRv6 Service Information Sub-TLV (33 bytes)
            Type: 1 Len: 30 Rsvd1: 0x0
            SRv6 SID: 2001:db8:aaaa:102::
            SID Flags: 0x0 Endpoint Behavior: 0x12 Rsvd2: 0x0
            SRv6 SID Sub-Sub-TLV
               Type: 1 Len: 6
               BL:48 NL:16 FL:20 AL:0 TL:20 TO:64
"
---snip---
```

PE-1 receives from BGP peer RR-3 (peer router id 2.2.2.3) the information for network prefix 172.16.222.1/32 that PE-2 (originator id 2.2.2.2) advertised, as displayed in the RIB In Entries section in the following example. PE-1 programs route prefix 172.16.222.1/32 in its local VPRN service route table and FIB. The presence of the SRv6 Services TLV indicates that the next hop is the VPRN SRv6 End.DT4 SID which, in turn, is resolved to the remote locator for PE-2. PE-2 expects the data with VPN label 524288. PE-2 has concatenated the hexadecimal value 0x80000 of this VPN label to the remote SRv6 SID prefix 2001:db8:aaaa:102:: to form the remote SRv6 full SID 2001:db8:aaaa:102:8000:: that PE-1 must use. PE-1 uses the path that corresponds with this information (flags field). PE-1 sends SRv6 encapsulated IPv4 data from the VPRN in an SRv6 tunnel to the remote locator prefix of PE-2 on its “int-PE-1-PE-2” interface (as is shown in the output of the **show router tunnel-table ipv6** command). PE-1 uses the VPRN 2 route table for the prefix lookup (VPRN imported field).

PE-1 advertises to BGP peer RR-3 (peer router id 2.2.2.3) the information for network prefix 172.16.211.1/32, as displayed in the RIB Out Entries section in the following example. RR-3 forwards this information to its BGP neighbors, in this case PE-2. PE-2 acts in a similar way as PE-1.

The following output shows the corresponding VPN-IPv4 BGP routes on PE-1:

```
[/]
A:admin@PE-1# show router bgp routes vpn-ipv4 hunt
===============================================================================
 BGP Router ID:2.2.2.1          AS:64500       Local AS:64500
===============================================================================
---snip---
===============================================================================
BGP VPN-IPv4 Routes
===============================================================================
-------------------------------------------------------------------------------
RIB In Entries
-------------------------------------------------------------------------------
```

```
Network        : 172.16.222.1/32
Nexthop        : 2001:db8::2:2
Route Dist.    : 192.0.2.2:2            VPN Label      : 524288
Path Id        : None
From           : 2001:db8::2:3
Res. Nexthop   : n/a
Local Pref.    : 100                    Interface Name : int-PE-1-PE-2
Aggregator AS  : None                   Aggregator     : None
Atomic Aggr.   : Not Atomic             MED            : None
AIGP Metric    : None                   IGP Cost       : 10
Connector      : None
Community      : target:64506:2
Cluster        : 3.3.3.3
Originator Id  : 2.2.2.2                Peer Router Id : 2.2.2.3
Fwd Class      : None                   Priority       : None
Flags          : Used Valid Best IGP
Route Source   : Internal
AS-Path        : No As-Path
Route Tag      : 0
Neighbor-AS    : n/a
Orig Validation: N/A
Source Class   : 0                      Dest Class     : 0
Add Paths Send : Default
Last Modified  : 00h00m33s
SRv6 TLV Type  : SRv6 L3 Service TLV (5)
SRv6 SubTLV    : SRv6 SID Information (1)
Sid            : 2001:db8:aaaa:102::
Full Sid       : 2001:db8:aaaa:102:8000::
Behavior       : End.DT4 (19)
SRv6 SubSubTLV : SRv6 SID Structure (1)
Loc-Block-Len  : 48                     Loc-Node-Len   : 16
Func-Len       : 20                     Arg-Len        : 0
Tpose-Len      : 20                     Tpose-offset   : 64
VPRN Imported  :  2

-------------------------------------------------------------------------------
RIB Out Entries
-------------------------------------------------------------------------------
Network        : 172.16.211.1/32
Nexthop        : 2001:db8::2:1
Route Dist.    : 192.0.2.1:2            VPN Label      : 524288
Path Id        : None
To             : 2001:db8::2:3
Res. Nexthop   : n/a
Local Pref.    : 100                    Interface Name : NotAvailable
Aggregator AS  : None                   Aggregator     : None
Atomic Aggr.   : Not Atomic             MED            : None
AIGP Metric    : None                   IGP Cost       : n/a
Connector      : None
Community      : target:64506:2
Cluster        : No Cluster Members
Originator Id  : None                   Peer Router Id : 2.2.2.3
Origin         : IGP
AS-Path        : No As-Path
Route Tag      : 0
Neighbor-AS    : n/a
Orig Validation: N/A
Source Class   : 0                      Dest Class     : 0
SRv6 TLV Type  : SRv6 L3 Service TLV (5)
SRv6 SubTLV    : SRv6 SID Information (1)
Sid            : 2001:db8:aaaa:101::
Full Sid       : 2001:db8:aaaa:101:8000::
Behavior       : End.DT4 (19)
SRv6 SubSubTLV : SRv6 SID Structure (1)
```

```
Loc-Block-Len  : 48                    Loc-Node-Len   : 16
Func-Len       : 20                    Arg-Len        : 0
Tpose-Len      : 20                    Tpose-offset   : 64

-------------------------------------------------------------------------------
Routes : 2
===============================================================================
```

For IPv6 data transport, VPRN End.DT6 behavior is needed. The IPv6 data transport uses a different VPN label 524287, resulting in a different full SRv6 SID ending with 7fff:f000::. PE-1 sends SRv6 encapsulated IPv6 data from the VPRN in an SRv6 tunnel to the remote locator prefix of PE-2 on its “int-PE-1-PE-2” interface (as is shown in the output of the **show router tunnel-table ipv6** command).

The following output shows the corresponding VPN-IPv6 BGP routes on PE-1:

```
[/]
A:admin@PE-1# show router bgp routes vpn-ipv6 hunt
===============================================================================
 BGP Router ID:2.2.2.1          AS:64500       Local AS:64500
===============================================================================
---snip---
===============================================================================
BGP VPN-IPv6 Routes
===============================================================================
-------------------------------------------------------------------------------
RIB In Entries
-------------------------------------------------------------------------------
Network        : 2001:db8:222::1/128
Nexthop        : 2001:db8::2:2
Route Dist.    : 192.0.2.2:2           VPN Label      : 524287
Path Id        : None
From           : 2001:db8::2:3
Res. Nexthop   : n/a
Local Pref.    : 100                   Interface Name : int-PE-1-PE-2
Aggregator AS  : None                  Aggregator     : None
Atomic Aggr.   : Not Atomic            MED            : None
AIGP Metric    : None                  IGP Cost       : 10
Connector      : None
Community      : target:64506:2
Cluster        : 3.3.3.3
Originator Id  : 2.2.2.2               Peer Router Id : 2.2.2.3
Fwd Class      : None                  Priority       : None
Flags          : Used Valid Best IGP
Route Source   : Internal
AS-Path        : No As-Path
Route Tag      : 0
Neighbor-AS    : n/a
Orig Validation: N/A
Source Class   : 0                     Dest Class     : 0
Add Paths Send : Default
Last Modified  : 00h00m33s
SRv6 TLV Type  : SRv6 L3 Service TLV (5)
SRv6 SubTLV    : SRv6 SID Information (1)
Sid            : 2001:db8:aaaa:102::
Full Sid       : 2001:db8:aaaa:102:7fff:f000::
Behavior       : End.DT6 (18)
SRv6 SubSubTLV : SRv6 SID Structure (1)
Loc-Block-Len  : 48                    Loc-Node-Len   : 16
Func-Len       : 20                    Arg-Len        : 0
Tpose-Len      : 20                    Tpose-offset   : 64
VPRN Imported  :  2

-------------------------------------------------------------------------------
```

```
RIB Out Entries
------------------------------------------------------------------------------
Network        : 2001:db8:211::1/128
Nexthop        : 2001:db8::2:1
Route Dist.    : 192.0.2.1:2           VPN Label      : 524287
Path Id        : None
To             : 2001:db8::2:3
Res. Nexthop   : n/a
Local Pref.    : 100                   Interface Name : NotAvailable
Aggregator AS  : None                  Aggregator     : None
Atomic Aggr.   : Not Atomic            MED            : None
AIGP Metric    : None                  IGP Cost       : n/a
Connector      : None
Community      : target:64506:2
Cluster        : No Cluster Members
Originator Id  : None                  Peer Router Id : 2.2.2.3
Origin         : IGP
AS-Path        : No As-Path
Route Tag      : 0
Neighbor-AS    : n/a
Orig Validation: N/A
Source Class   : 0                     Dest Class     : 0
SRv6 TLV Type  : SRv6 L3 Service TLV (5)
SRv6 SubTLV    : SRv6 SID Information (1)
Sid            : 2001:db8:aaaa:101::
Full Sid       : 2001:db8:aaaa:101:7fff:f000::
Behavior       : End.DT6 (18)
SRv6 SubSubTLV : SRv6 SID Structure (1)
Loc-Block-Len  : 48                    Loc-Node-Len   : 16
Func-Len       : 20                    Arg-Len        : 0
Tpose-Len      : 20                    Tpose-offset   : 64

------------------------------------------------------------------------------
Routes : 2
==============================================================================
```

Verify that there are additional local SRv6 SIDs for PE-1 and PE-2. These local SRv6 SIDs correspond with the additional SRv6 behavior that is configured on the locator for the data transport between the local and the remote VPRN. Because RR-3 does not have SRv6 configuration, RR-3 does not have local SRv6 SIDs.

```
[/]
A:admin@PE-2# show router segment-routing-v6 local-sid

==============================================================================
Segment Routing v6 Local SIDs
==============================================================================
SID                                                 Type           Function
  Locator
  Context
------------------------------------------------------------------------------
---snip---
2001:db8:aaaa:102:7fff:f000::                       End.DT6        524287
  PE-2_loc
  SvcId: 2 Name: VPRN_2
2001:db8:aaaa:102:8000::                            End.DT4        524288
  PE-2_loc
  SvcId: 2 Name: VPRN_2
------------------------------------------------------------------------------
SIDs : 5
------------------------------------------------------------------------------
==============================================================================
```

Verify that there is SRv6 information for the VPRN service with service id 2.

```
[/]
A:admin@PE-2# show service id 2 segment-routing-v6 detail

===============================================================================
Segment Routing v6 Instance 1 Service 2
===============================================================================
Locator
  Type         Function  SID                                          Status
-------------------------------------------------------------------------------
PE-2_loc
  End.DT4        *524288 2001:db8:aaaa:102:8000::                     ok
  End.DT6        *524287 2001:db8:aaaa:102:7fff:f000::                ok
===============================================================================
Legend: * - System allocated
```

At this point, verify that data transport is possible between the local VPRN on PE-1 and the remote VPRN on PE-2.

```
[/]
A:admin@PE-1# ping 172.16.222.1 router-instance "VPRN_2"
PING 172.16.222.1 56 data bytes
64 bytes from 172.16.222.1: icmp_seq=1 ttl=64 time=1.53ms.
---snip---
---- 172.16.222.1 PING Statistics ----
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min = 1.48ms, avg = 1.53ms, max = 1.59ms, stddev = 0.035ms

[/]
A:admin@PE-1# ping 2001:db8:222::1 router-instance "VPRN_2"
PING 2001:db8:222::1 56 data bytes
64 bytes from 2001:db8:222::1 icmp_seq=1 hlim=64 time=1.29ms.
---snip---
---- 2001:db8:222::1 PING Statistics ----
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min = 1.28ms, avg = 1.43ms, max = 1.78ms, stddev = 0.181ms

[/]
A:admin@PE-1# traceroute 172.16.222.1 router-instance "VPRN_2"
traceroute to 172.16.222.1, 30 hops max, 40 byte packets
  1  172.16.222.1 (172.16.222.1)    1.61 ms  1.49 ms  1.73 ms

[/]
A:admin@PE-1# traceroute 2001:db8:222::1 router-instance "VPRN_2"
traceroute to 2001:db8:222::1, 30 hops max, 60 byte packets
  1  2001:db8:222::1 (2001:db8:222::1)    1.36 ms  1.47 ms  1.57 ms
```

The result of the verification complies with the route table for the local VPRN on PE-1, which now also contains routes for the loopback addresses in the remote VPRN on PE-2. The same is true for data transport between the remote VPRN on PE-2 and the local VPRN on PE-1.

```
A:admin@PE-1# show router 2 route-table ipv4

===============================================================================
Route Table (Service: 2)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                         Active     Metric
-------------------------------------------------------------------------------
172.16.211.1/32                               Local   Local     00h08m40s  0
```

```
       lb_itf_vprn                                        Y            0
172.16.222.1/32                                 Remote  BGP VPN   00h01m00s  170
       2001:db8:aaaa:102:8000:: (tunneled:SRV6)            Y            10
-------------------------------------------------------------------------------
No. of Routes: 2
---snip---
===============================================================================
```

```
[/]
A:admin@PE-1# show router 2 route-table ipv6 all

===============================================================================
IPv6 Route Table (Service: 2)
===============================================================================
Dest Prefix[Flags]                              Type    Proto     Age        Pref
      Next Hop[Interface Name]                           Active     Metric
-------------------------------------------------------------------------------
2001:db8:211::1/128                             Local   Local     00h08m38s  0
       lb_itf_vprn                                        Y            0
2001:db8:222::1/128                             Remote  BGP VPN   00h01m00s  170
       2001:db8:aaaa:102:7fff:f000:: (tunneled:SRV6)       Y            10
-------------------------------------------------------------------------------
No. of Routes: 2
---snip---
===============================================================================
```

Conclusion

SRv6 shortest path routing can be used as an IPv6 transport for implementing VPRN services across an IPv6 service provider network.

Segment Routing with IS-IS Control Plane

This chapter provides information about Segment Routing (SR) with Intermediate System to Intermediate System (IS-IS) control plane.

Topics in this chapter include:

Applicability

Segment routing is supported in SR OS Release 13.0, and later. This chapter was initially written for SR OS Release 13.0.R3, but the MD-CLI in the current edition corresponds to SR OS Release 25.10.R1.

Overview

Segment Routing (SR) is a technology for IP/Multi-Protocol Label Switching (MPLS) networks that enables source routing. With source routing, operators can specify a forwarding path, from ingress to egress, that is independent of the shortest path determined by the Interior Gateway Protocol (IGP).

The main benefit of segment routing compared to other source routing protocols (such as ReSource reserVation Protocol with Traffic Engineering (RSVP-TE)) is that, from a control plane perspective, no signaling protocol is required. Segment routing provides a path or tunnel, encoded as a sequential list of sub-paths or segments that are advertised within the segment routing domain, using extensions to well-known link state routing protocols, such as IS-IS or Open Shortest Path First (OSPF).

Implementation

A segment routing tunnel can contain a single segment that represents the destination node, or it can contain a list of segments that the tunnel must traverse. The tunnel can be established over an IPv4/IPv6 MPLS or IPv6 data plane, encoded as a stack of MPLS labels or as a number of IPv6 addresses contained in an IPv6 extension header.

Network elements are modeled as segments. For each segment, IGP advertises an identifier referred to as a segment ID (SID).

The two segment types are:

- *Prefix segment* — Globally unique and allocated from a Segment Routing Global Block (SRGB), typically multi-hop and signaled by the IGP. It is the Equal Cost Multi-Path ECMP-aware shortest path IGP route to a related prefix. A typical example of a prefix segment is a node SID. Within the SR OS implementation, the node SID is either the system address or another interface address in the Global

Routing Table (GRT) of type loopback. Node SIDs are advertised in IS-IS using a prefix SID sub-TLV (Type Length Value).

- *Adjacency segment* — Locally unique and allocated from the local dynamic label space, so that other routers in the SR domain can use the same label space. Adjacency segments are signaled by the IGP. Within the SR OS implementation, adjacency SIDs are automatically assigned and advertised when the SR context within the IGP instance is enabled. Adjacency SIDs are advertised in IS-IS using an adjacency SID sub-TLV.

To make prefix segments globally unique within the segment routing domain, an indexing mechanism is required, because production networks consist of multiple vendors and multiple products. As a result, it is often difficult to agree on a common SRGB for the prefix SIDs.

All routers within the SR domain are expected to configure and advertise the same Prefix SID index range for an IGP instance. The label value used by each router to represent a prefix can be local to that router by the use of an offset label, referred to as a start label:

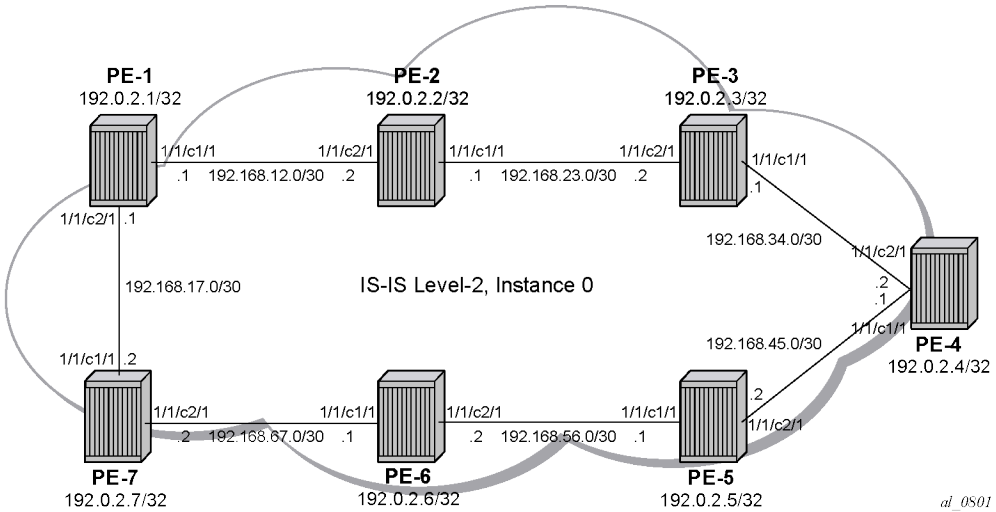
Local label (for a prefix) = local start label + {Prefix SID index}

Within the SR OS implementation, prefix Loop-Free Alternate (LFA) is supported for segment routing to improve the Fast ReRoute (FRR) coverage. Remote LFA (RLFA) is also supported. With RLFA, segment routing shortest path tunnels are used as a virtual LFA or repair tunnel toward the PQ node. RLFA is described in Remote Loop-Free Alternate Node Protection. Topology-independent LFA (TI-LFA) is also supported and is described in Topology-Independent Loop-Free Alternate for Link Protection.

Configuration

The following example uses IS-IS as an IGP protocol, with an MPLS data plane and services enabled using LFA and RLFA. Figure 56: Example topology shows the example topology with seven PEs.

Figure 56: Example topology



Initial configuration

The system and IP interface addresses are configured according to Figure 56: Example topology.

IS-IS level 2 is selected as the IGP to distribute routing information between all PEs. All IS-IS interfaces are of type point-to-point to avoid running the Designated Router/Backup Designated Router (DR/BDR) election process.

Segment routing configuration

Before enabling segment routing on a router, define a dedicated SRGB. This SRGB is required on each individual router part of the SR domain and is used to allocate the Prefix SIDs.

By default, an SRGB is not instantiated and, when configured by the operator, it is taken from the system dynamic label range. By default, the following label ranges are available:

```
[/]
A:admin@PE-1# show router mpls-labels label-range

===============================================================================
Label Ranges
===============================================================================
Label Type      Start Label End Label   Aging       Available   Total
-------------------------------------------------------------------------------
Static          32          18431       -           18400       18400
Dynamic         18432       524287      0           505856      505856
    Seg-Route   0           0           -           0           0
===============================================================================
```

For simplicity, the same SRGB is used in this example for all SR domain routers. Within the command, a start value and end value define the size of the SRGB. The following command configures an SRGB of 10000 MPLS labels, from label 20000 to label 29999:

```
# on PE-1, PE-2, PE-3, PE-4, PE-5, PE-6, PE-7:
configure {
    router "Base" {
        mpls-labels {
            sr-labels {
                start 20000
                end 29999
            }
        }
```

```
[/]
A:admin@PE-1# show router mpls-labels label-range

===============================================================================
Label Ranges
===============================================================================
Label Type      Start Label End Label   Aging       Available   Total
-------------------------------------------------------------------------------
Static          32          18431       -           18400       18400
Dynamic         18432       524287      0           505756      505856
    Seg-Route   20000       29999       -           0           10000
===============================================================================
```

This command is repeated for all other nodes. The allocated MPLS labels are only for the prefix SIDs. The adjacency SIDs, which are only locally unique, are taken from the dynamic range; in this example, between 18432 and 524287.

The following steps are taken to configure segment routing:

1. Enable router capability in IGP instance
2. Define prefix SID index range
3. Configure node SID
4. Enable segment routing

Enable router capability in IGP instance

It is mandatory to enable the router-capability parameter inside the IS-IS instance, to advertise SR support among the IS-IS adjacencies. By configuring this command within the IGP instance, the SR capability sub-TLV is propagated and is used to indicate the index range and the start label. The SR algorithm sub-TLV is also used to advertise the algorithm used for path calculations. Only Shortest Path First (SPF) (value 0) is defined. This is configured as follows:

```
# on PE-1, PE-2, PE-3, PE-4, PE-5, PE-6, PE-7:
configure {
    router "Base" {
        isis 0 {
            advertise-router-capability area
```

The flooding parameter is a mandatory parameter in this CLI command. The keyword **area** or **as** indicates that the router capabilities label switched path (LSP) should be advertised throughout the same level or throughout the whole Autonomous System (AS). In the preceding example, all routers belong to the same level, so the **area** argument is sufficient. When the SR context within the IGP instance is enabled, both IS-IS sub-TLVs are flooded.

Define prefix SID index range

The SR OS implementation for SR provides two mutually exclusive modes of operation to define the Prefix SID index range: global mode and per-instance mode. Per-instance mode is useful in a seamless MPLS environment when multiple IGP instances are used. The main difference between the modes is the way that the start label and index range are calculated.

Table 8: Mode comparison compares global mode and per-instance mode:

Table 8: Mode comparison

Global mode	Per-instance mode
Applicable for all IGP instances on that node	Applicable for one dedicated IGP instance
Start label is first label of SRGB	Start label is configurable (but part of SRGB range); use of non-overlapping sub-ranges of SRGB

Global mode	Per-instance mode
Prefix SID index range is "size" of SRGB	Prefix SID index-range is configurable
If SRGB needs to change, disable SR and delete prefix-SID-ranges in all IGP instances	If prefix SID index and/or label range needs to change, disable SR in that specific IGP instance
SW checks whether any allocated SID index/label goes out of range. SW checks also for overlaps of the resulting net label value range across IGP instances.	

For simplicity, global mode is used for this example, as follows:

```
# on PE-1, PE-2, PE-3, PE-4, PE-5, PE-6, PE-7:
configure {
    router "Base" {
        isis 0 {
            segment-routing {
                prefix-sid-range {
                    global
                }
```

Configure node SID

To be able to set up SR shortest path tunnels to all routers of the SR domain, each router needs to be uniquely defined within the SR domain. Therefore, the system address or other loopback interface in the GRT is assigned a node SID.

Both an IPv4 and an IPv6 node SID can be configured using an index or a label. The node SID for a node has a label equal to the sum of the start label (20000), which is the first label of the SRGB on all nodes, and the index on the node.

It is possible to configure IS-IS node SIDs, protocol-independent node SIDs, or both:

- IS-IS prefix SIDs
- Protocol-independent prefix SIDs

IS-IS prefix SIDs

IS-IS prefix SIDs are configured on the system interface in the IS-IS context. The following IS-IS IPv4 node SIDs are applicable for IS-IS instance 0 and cannot be shared with other IGPs. In this example, an IS-IS IPv4 node SID is configured using an index on all PEs except PE-7, as follows:

```
# on PE-1:
configure {
    router "Base" {
        isis 0 {
            interface "system" {
                ipv4-node-sid {
                    index 1  # on PE-2: index 2, on PE-3: index 3, ...
```

Because the SRGB is the same on all nodes, each node in the network can be reached using the same MPLS label in the range from 20000 to 29999. For example, the node SID for PE-1 on all nodes is 20001 (= start label 20000 + index 1).

When there is one consistent SRGB for the SR domain, the SR OS allows the use of absolute MPLS label values instead of index values. For example, on PE-1, an operator can use an explicit MPLS label value, as follows:

```
# on PE-1:
configure {
    router "Base" {
        isis 0 {
            interface "system" {
                ipv4-node-sid {
                    label 20001
```

Internally, this explicit value is translated into an index value (index-value 1) before advertising it toward its neighbors, taking into account the prefix SID index-range mode (global or per-instance) and the SRGB.

Protocol-independent prefix SIDs

In this example, the configuration on PE-7 does not include an IS-IS prefix SID; the node SID is protocol-independent instead. This node SID can be used in multiple IGP instances.

Protocol-independent prefix SIDs are configured in the **configure router segment-routing sr-mpls** context, as follows:

```
# on PE-7:
configure {
    router "Base" {
        segment-routing {
            sr-mpls {
                prefix-sids "system"
                    node-sid true
                    ipv4-sid {
                        index 4007        # protocol-independent SID shared by IGPs
                    }
```

With one consistent SRGB for the SR domain, absolute values can be configured instead, as follows:

```
# on PE-1:
configure {
    router "Base" {
        segment-routing {
            sr-mpls {
                prefix-sids "system"
                    node-sid true
                    ipv4-sid {
                        label 24007        # protocol-independent SID shared by IGPs
                    }
```

When both an IS-IS node SID and a protocol-independent node SID are configured on the same node, the IS-IS node SID is advertised by IS-IS, not the protocol-independent node SID.

Enable segment routing

Enable SR context within the IGP instance, as follows:

```
# on PE-1, PE-2, PE-3, PE-4, PE-5, PE-6, PE-7:
configure {
    router "Base" {
        isis 0 {
            segment-routing {
                admin-state enable
```

Verification

After enabling the SR context within an IGP instance, the SR capability sub-TLV, and the SR algorithm sub-TLV between all routers within the SR domain, are flooded. The following show command displays the SR related router capability information on PE-1:

```
[/]
A:admin@PE-1# show router isis capabilities level 2

===============================================================================
Rtr Base ISIS Instance 0 Capabilities
===============================================================================

Displaying Level 2 capabilities
-------------------------------------------------------------------------------
LSP ID    : PE-1.00-00
  Router Cap : 192.0.2.1, D:0, S:0
    TE Node Cap : B E M  P
    SR Cap: IPv4 MPLS-IPv6
       SRGB Base:20000, Range:10000
    SR Alg: metric based SPF
    Node MSD Cap: BMI : 12 ERLD : 15

LSP ID    : PE-2.00-00
  Router Cap : 192.0.2.2, D:0, S:0
    TE Node Cap : B E M  P
    SR Cap: IPv4 MPLS-IPv6
       SRGB Base:20000, Range:10000
    SR Alg: metric based SPF
    Node MSD Cap: BMI : 12 ERLD : 15

LSP ID    : PE-3.00-00
  Router Cap : 192.0.2.3, D:0, S:0
    TE Node Cap : B E M  P
    SR Cap: IPv4 MPLS-IPv6
       SRGB Base:20000, Range:10000
    SR Alg: metric based SPF
    Node MSD Cap: BMI : 12 ERLD : 15

LSP ID    : PE-4.00-00
  Router Cap : 192.0.2.4, D:0, S:0
    TE Node Cap : B E M  P
    SR Cap: IPv4 MPLS-IPv6
       SRGB Base:20000, Range:10000
    SR Alg: metric based SPF
    Node MSD Cap: BMI : 12 ERLD : 15
```

```
LSP ID    : PE-5.00-00
  Router Cap : 192.0.2.5, D:0, S:0
    TE Node Cap : B E M  P
    SR Cap: IPv4 MPLS-IPv6
       SRGB Base:20000, Range:10000
    SR Alg: metric based SPF
    Node MSD Cap: BMI : 12 ERLD : 15

LSP ID    : PE-6.00-00
  Router Cap : 192.0.2.6, D:0, S:0
    TE Node Cap : B E M  P
    SR Cap: IPv4 MPLS-IPv6
       SRGB Base:20000, Range:10000
    SR Alg: metric based SPF
    Node MSD Cap: BMI : 12 ERLD : 15

LSP ID    : PE-7.00-00
  Router Cap : 192.0.2.7, D:0, S:0
    TE Node Cap : B E M  P
    SR Cap: IPv4 MPLS-IPv6
       SRGB Base:20000, Range:10000
    SR Alg: metric based SPF
    Node MSD Cap: BMI : 12 ERLD : 15

Level (2) Capability Count : 7
===============================================================================
```

A similar output occurs for each router in the SR domain.

After enabling the SR context within the IGP instance, the assigned index for each locally configured prefix SID is advertised. After the advertisement of prefix SIDs, MPLS data plane Ingress Label Mapping (ILM) is programmed with a pop operation. In this context, a show command can be used to display the prefix SIDs, in order, within the SR domain, for example, on PE-1:

```
[/]
A:admin@PE-1# show router isis prefix-sids

===============================================================================
Rtr Base ISIS Instance 0 Prefix/SID Table
===============================================================================
Prefix                          SID        Lvl/Typ    SRMS   AdvRtr
                                 Shared                 MT     Flags
-------------------------------------------------------------------------------
192.0.2.1/32                    1          2/Int.      N     PE-1
                                    No                    0      NnP
192.0.2.2/32                    2          2/Int.      N     PE-2
                                    N.A.                  0      NnP
192.0.2.3/32                    3          2/Int.      N     PE-3
                                    N.A.                  0      NnP
192.0.2.4/32                    4          2/Int.      N     PE-4
                                    N.A.                  0      NnP
192.0.2.5/32                    5          2/Int.      N     PE-5
                                    N.A.                  0      NnP
192.0.2.6/32                    6          2/Int.      N     PE-6
                                    N.A.                  0      NnP
192.0.2.7/32                    4007       2/Int.      N     PE-7
                                    N.A.                  0      NnP
-------------------------------------------------------------------------------
No. of Prefix/SIDs: 7 (7 unique)
-------------------------------------------------------------------------------
SRMS:   Y/N  = prefix SID advertised by SR Mapping Server (Y) or not (N)
        S    = SRMS prefix SID is selected to be programmed
Flags:  R    = Re-advertisement
```

```
        N    = Node-SID
        nP   = no penultimate hop POP
        E    = Explicit-Null
        V    = Prefix-SID carries a value
        L    = value/index has local significance
Shared: Yes  = local shared Node-SID
        No   = not a local shared Node-SID
        N.A. = not applicable for Remote prefix-sid
===============================================================================
```

By default, the SR OS implementation sets the node SID (or *N*–flag) and no Penultimate hop PoP (or *nP*–flag) inside the prefix SID TLV. Another useful flag that can be set is the re-advertisement (or R-flag). The R-flag is set when a prefix SID is propagated between levels or areas, or redistribution is in place (from another protocol).

De node SID of PE-1 is 20001 is an IS-IS node SID, so it is not shared with other IGPs; only the node SID of PE-7 is a protocol-independent node SID that can be shared with other IGPs, but that is only applicable on PE-7 itself, as follows:

```
[/]
A:admin@PE-7# show router isis prefix-sids

===============================================================================
Rtr Base ISIS Instance 0 Prefix/SID Table
===============================================================================
Prefix                          SID        Lvl/Typ    SRMS   AdvRtr
                                 Shared                 MT     Flags
-------------------------------------------------------------------------------
192.0.2.1/32                    1          2/Int.      N     PE-1
                                    N.A.                   0      NnP
192.0.2.2/32                    2          2/Int.      N     PE-2
                                    N.A.                   0      NnP
192.0.2.3/32                    3          2/Int.      N     PE-3
                                    N.A.                   0      NnP
192.0.2.4/32                    4          2/Int.      N     PE-4
                                    N.A.                   0      NnP
192.0.2.5/32                    5          2/Int.      N     PE-5
                                    N.A.                   0      NnP
192.0.2.6/32                    6          2/Int.      N     PE-6
                                    N.A.                   0      NnP
192.0.2.7/32                    4007       2/Int.      N     PE-7
                                    Yes                    0      NnP
-------------------------------------------------------------------------------
No. of Prefix/SIDs: 7 (7 unique)
-------------------------------------------------------------------------------
SRMS:   Y/N  = prefix SID advertised by SR Mapping Server (Y) or not (N)
        S    = SRMS prefix SID is selected to be programmed
Flags:  R    = Re-advertisement
        N    = Node-SID
        nP   = no penultimate hop POP
        E    = Explicit-Null
        V    = Prefix-SID carries a value
        L    = value/index has local significance
Shared: Yes  = local shared Node-SID
        No   = not a local shared Node-SID
        N.A. = not applicable for Remote prefix-sid
===============================================================================
```

Prefix SID information can also be viewed within the IGP database attached to (extended) IP prefix reachability TLVs. For example, on PE-1, as follows:

```
[[/]
A:admin@PE-1# show router isis database PE-1.00-00 detail level 2

===============================================================================
Rtr Base ISIS Instance 0 Database (detail)
===============================================================================

Displaying Level 2 database
-------------------------------------------------------------------------------
LSP ID    : PE-1.00-00                                Level     : L2
Sequence  : 0x6                  Checksum  : 0xf93a   Lifetime  : 1106
Version   : 1                    Pkt Type  : 20       Pkt Ver   : 1
Attributes: L1L2                 Max Area  : 3        Alloc Len : 1492
SYS ID    : 1920.0000.2001       SysID Len : 6        Used Len  : 254

TLVs :
  Supp Protocols:
    Protocols     : IPv4
  IS-Hostname   : PE-1
  Router ID   :
    Router ID   : 192.0.2.1
  Router Cap : 192.0.2.1, D:0, S:0
    TE Node Cap : B E M  P
    SR Cap: IPv4 MPLS-IPv6
       SRGB Base:20000, Range:10000
    SR Alg: metric based SPF
    Node MSD Cap: BMI : 12 ERLD : 15
---snip---

  Internal Reach:
    ---snip---
    Default Metric: (I) 0
    Delay Metric  : (I) 0
    Expense Metric: (I) 0
    Error Metric  : (I) 0
    IP Address    : 192.0.2.1
    IP Mask       : 255.255.255.255
  I/F Addresses :
    I/F Address   : 192.0.2.1
    ---snip---

  TE IP Reach   :
    ---snip---
    Default Metric  : 0
    Control Info:   S, prefLen 32
    Prefix   : 192.0.2.1
    Sub TLV   :
      Prefix-SID Index:1, Algo:0, Flags:NnP

Level (2) LSP Count : 1
-------------------------------------------------------------------------------
---snip---
Prefix-SID Flags : R = Re-advertisement Flag
                   N = Node-SID Flag
                   nP = no penultimate hop POP
                   E = Explicit-Null Flag
                   V = Prefix-SID carries a value
                   L = value/index has local significance
---snip---
```

```
===============================================================================
```

After enabling the SR context within the IGP instance, adjacency SIDs are also automatically assigned and advertised for each formed adjacency over an IP interface. From a data plane perspective, one local adjacency SID consumes one ILM entry, programming a pop operation.

Similar to prefix SIDs, adjacency SID information can be viewed within the IGP database attached to IS neighbor TLVs, as follows:

```
[/]
A:admin@PE-1# show router isis database PE-1.00-00 detail level 2 | match Adj pre-lines 6 post-
lines 1
  TE IS Nbrs   :
    Nbr   : PE-2.00
    Default Metric  : 10
    Sub TLV Len     : 19
    IF Addr   : 192.168.12.1
    Nbr IP    : 192.168.12.2
    Adj-SID: Flags:v4VL Weight:0 Label:524287
  TE IS Nbrs   :
    Nbr   : PE-7.00
    Default Metric  : 10
    Sub TLV Len     : 19
    IF Addr   : 192.168.17.1
    Nbr IP    : 192.168.17.2
    Adj-SID: Flags:v4VL Weight:0 Label:524286
  TE IP Reach   :

---snip---
Adj-SID Flags    : v4/v6 = IPv4 or IPv6 Address-Family
                   B = Backup Flag
                   V = Adj-SID carries a value
                   L = value/index has local significance
                   S = Set of Adjacencies
                   P = Persistently allocated
---snip---
```

By default, the SR OS implementation sets the value (*V*-flag), meaning that the adjacency SID carries a value (as opposed to an index). Also, the local *L*-flag is set by default, meaning that the adjacency SID has only local significance. The *v4*-flag set to 0 means that the adjacency SID references to an adjacency with outgoing IPv4 encapsulation.

Another way to display adjacency SID information is using the **show router isis adjacency detail** command.

```
[/]
A:admin@PE-1# show router isis adjacency "int-PE-1-PE-2" detail

===============================================================================
Rtr Base ISIS Instance 0 Adjacency (detail)
===============================================================================
Hostname    : PE-2
SystemID    : 1920.0000.2002                  SNPA        : 02:0e:01:01:00:0b
Interface   : int-PE-1-PE-2                   Up Time     : 0d 00:05:31
State       : Up                              Priority    : 0
Nbr Sys Typ : L2                              L. Circ Typ : L2
Hold Time   : 27                              Max Hold    : 27
Adj Level   : L2                              MT Enabled  : No
Topology    : Unicast

IPv6 Neighbor         : ::
```

```
IPv4 Neighbor         : 192.168.12.2
IPv4 Adj SID          : Label 524287
IPv4 SID Protect      : No
Restart Support       : Disabled
Restart Status        : Not currently being helped
Restart Supressed     : Disabled
Number of Restarts    : 0
Last Restart at       : Never

===============================================================================
```

```
[/]
A:admin@PE-1# show router isis adjacency "int-PE-1-PE-7" detail

===============================================================================
Rtr Base ISIS Instance 0 Adjacency (detail)
===============================================================================
Hostname    : PE-7
SystemID    : 1920.0000.2007                  SNPA        : 02:24:01:01:00:01
Interface   : int-PE-1-PE-7                   Up Time     : 0d 00:04:55
State       : Up                              Priority    : 0
Nbr Sys Typ : L2                              L. Circ Typ : L2
Hold Time   : 27                              Max Hold    : 27
Adj Level   : L2                              MT Enabled  : No
Topology    : Unicast

IPv6 Neighbor         : ::
IPv4 Neighbor         : 192.168.17.2
IPv4 Adj SID          : Label 524286
IPv4 SID Protect      : No
Restart Support       : Disabled
Restart Status        : Not currently being helped
Restart Supressed     : Disabled
Number of Restarts    : 0
Last Restart at       : Never

===============================================================================
```

Finally, when enabling the SR context within the IGP instance, the SR module resolves received prefixes with prefix SID sub-TLVs present. As a result, MPLS data plane resources are consumed. The ILM is programmed with a swap operation and the label-to-next-hop-label-forwarding-entry (LTN) with a push operation, both pointing to the primary and/or LFA next-hop label forwarding entry (NHLFE). Also, an SR tunnel is added in the Tunnel Table Manager (TTM). As a result, an SR shortest path tunnel is set up to each other router that is part of the SR domain. Now, SR shortest path tunnels can be used for all users of TTM.

Example 1 - VPRN service with LFA and RLFA enabled

In the network topology of Figure 56: Example topology, no LDP and RSVP-TE signaling protocols are enabled. Each router of the SR domain has a full mesh of SR shortest path tunnels to the other routers, and no LDP and RSVP-TE LSPs are present. For example, on PE-1, the TTM looks as follows:

```
[/]
A:admin@PE-1# show router tunnel-table

===============================================================================
IPv4 Tunnel Table (Router: Base)
===============================================================================
```

```
Destination           Owner     Encap TunnelId  Pref   Nexthop        Metric
   Color
-------------------------------------------------------------------------------
192.0.2.2/32          isis (0)  MPLS  524291    11     192.168.12.2   10
192.0.2.3/32          isis (0)  MPLS  524292    11     192.168.12.2   20
192.0.2.4/32          isis (0)  MPLS  524293    11     192.168.12.2   30
192.0.2.5/32          isis (0)  MPLS  524295    11     192.168.17.2   30
192.0.2.6/32          isis (0)  MPLS  524296    11     192.168.17.2   20
192.0.2.7/32          isis (0)  MPLS  524294    11     192.168.17.2   10
192.168.12.2/32       isis (0)  MPLS  524289    11     192.168.12.2   0
192.168.17.2/32       isis (0)  MPLS  524290    11     192.168.17.2   0
-------------------------------------------------------------------------------
Flags: B = BGP or MPLS backup hop available
       L = Loop-Free Alternate (LFA) hop available
       E = Inactive best-external BGP route
       k = RIB-API or Forwarding Policy backup hop
===============================================================================
```

The objective is to configure a VPRN between PE-1 and PE-7, using SR shortest path tunnels as transport tunnel, as follows:

```
# on PE-1:
configure {
    service {
        vprn "VPRN-100" {
            admin-state enable
            service-id 100
            customer "1"
            autonomous-system 64496
            bgp-ipvpn {
                mpls {
                    admin-state enable
                    route-distinguisher "192.0.2.1:100"
                    vrf-target {
                        community "target:64496:100"
                    }
                    auto-bind-tunnel {
                        resolution any
                    }
                }
            }
            interface "loopback" {
                loopback true
                ipv4 {
                    primary {
                        address 10.10.1.1
                        prefix-length 32
                    }
                }
            }
```

```
# on PE-7:
configure {
    service {
        vprn "VPRN-100" {
            admin-state enable
            service-id 100
            customer "1"
            autonomous-system 64496
            bgp-ipvpn {
                mpls {
                    admin-state enable
```

```
                    route-distinguisher "192.0.2.7:100"
                    vrf-target {
                        community "target:64496:100"
                    }
                    auto-bind-tunnel {
                        resolution any
                    }
                }
            }
            interface "loopback" {
                loopback true
                ipv4 {
                    primary {
                        address 10.10.1.7
                        prefix-length 32
                    }
                }
            }
```

Within the VPRN service configuration, a loopback interface is created on both PEs to verify the transport mechanism. Tunnel information displaying the MPLS label value is retrieved using the **show router fp-tunnel-table <slot number>** command, as follows:

```
[/]
A:admin@PE-1# show router fp-tunnel-table 1 192.0.2.7/32

===============================================================================
IPv4 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                Protocol         Tunnel-ID
  Lbl/SID
    NextHop                                                 Intf/Tunnel
  Lbl/SID (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
192.0.2.7/32                               SR-ISIS-0         524291
  24007
    192.168.17.2                                           1/1/c2/1:1000
-------------------------------------------------------------------------------
Total Entries : 1
-------------------------------------------------------------------------------
===============================================================================
```

This means that, when traffic arrives on PE-1, the MPLS label 24007 is pushed to reach destination PE-7. Because, in this example, the prefix SID index range global mode is used, the value 24007 comes from the start label on PE-7 (first label of the SRGB, which is 20000, plus the configured index value of node SID PE-7 (4007)), so 24007.

Enabling prefix LFA within the IS-IS context on PE-1 enables LFA/FRR protection:

```
# on PE-1:
configure {
    router "Base" {
        isis 0 {
            loopfree-alternate {
```

The IS-IS LFA coverage on PE-1 is as follows:

```
[/]
A:admin@PE-1# show router isis lfa-coverage

===============================================================================
Rtr Base ISIS Instance 0 LFA Coverage
===============================================================================
Topology         Level   Node           IPv4                IPv6
-------------------------------------------------------------------------------
IPV4 Unicast     L1      0/0(0%)        3/11(27%)           0/0(0%)
IPV6 Unicast     L1      0/0(0%)        0/0(0%)             0/0(0%)
IPV4 Multicast   L1      0/0(0%)        0/0(0%)             0/0(0%)
IPV6 Multicast   L1      0/0(0%)        0/0(0%)             0/0(0%)
IPV4 Unicast     L2      2/6(33%)       3/11(27%)           0/0(0%)
IPV6 Unicast     L2      0/0(0%)        0/0(0%)             0/0(0%)
IPV4 Multicast   L2      0/0(0%)        0/0(0%)             0/0(0%)
IPV6 Multicast   L2      0/0(0%)        0/0(0%)             0/0(0%)
===============================================================================
```

Next-hop LFA protection is present for node PE-4, node PE-5, and the link between PE-4 and PE-5, as follows:

```
[/]
A:admin@PE-1# show router route-table alternative

===============================================================================
Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                              Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
      Alt-NextHop                                                 Alt-
                                                                  Metric
-------------------------------------------------------------------------------
192.0.2.1/32                                    Local   Local     00h16m16s  0
       system                                                         0
192.0.2.2/32                                    Remote  ISIS      00h11m57s  18
       192.168.12.2                                                   10
192.0.2.3/32                                    Remote  ISIS      00h11m54s  18
       192.168.12.2                                                   20
192.0.2.4/32                                    Remote  ISIS      00h11m54s  18
       192.168.12.2                                                   30
       192.168.17.2 (LFA)                                             40
192.0.2.5/32                                    Remote  ISIS      00h11m03s  18
       192.168.17.2                                                   30
       192.168.12.2 (LFA)                                             40
192.0.2.6/32                                    Remote  ISIS      00h11m03s  18
       192.168.17.2                                                   20
192.0.2.7/32                                    Remote  ISIS      00h11m03s  18
       192.168.17.2                                                   10
192.168.12.0/30                                 Local   Local     00h16m16s  0
       int-PE-1-PE-2                                                  0
192.168.17.0/30                                 Local   Local     00h16m16s  0
       int-PE-1-PE-7                                                  0
192.168.23.0/30                                 Remote  ISIS      00h15m57s  18
       192.168.12.2                                                   20
192.168.34.0/30                                 Remote  ISIS      00h15m47s  18
       192.168.12.2                                                   30
192.168.45.0/30                                 Remote  ISIS      00h14m35s  18
       192.168.12.2                                                   40
       192.168.17.2 (LFA)                                             50
192.168.56.0/30                                 Remote  ISIS      00h15m12s  18
       192.168.17.2                                                   30
```

```
192.168.67.0/30                                Remote  ISIS      00h15m12s  18
       192.168.17.2                                                    20
-------------------------------------------------------------------------------
No. of Routes: 14
Flags: n = Number of times nexthop is repeated
       Backup = BGP backup route
       LFA = Loop-Free Alternate nexthop
       S = Sticky ECMP requested
===============================================================================
```

```
[/]
A:admin@PE-1# show router fp-tunnel-table 1

===============================================================================
IPv4 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                Protocol         Tunnel-ID
  Lbl/SID
    NextHop                                                 Intf/Tunnel
  Lbl/SID (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
192.0.2.2/32                               SR-ISIS-0         524293
  20002
    192.168.12.2                                           1/1/c1/1:1000
192.0.2.3/32                               SR-ISIS-0         524294
  20003
    192.168.12.2                                           1/1/c1/1:1000
192.0.2.4/32                               SR-ISIS-0         524295
  20004
    192.168.12.2                                           1/1/c1/1:1000
  20004
    192.168.17.2(B)                                        1/1/c2/1:1000
192.0.2.5/32                               SR-ISIS-0         524296
  20005
    192.168.17.2                                           1/1/c2/1:1000
  20005
    192.168.12.2(B)                                        1/1/c1/1:1000
192.0.2.6/32                               SR-ISIS-0         524292
  20006
    192.168.17.2                                           1/1/c2/1:1000
192.0.2.7/32                               SR-ISIS-0         524291
  24007
    192.168.17.2                                           1/1/c2/1:1000
192.168.12.2/32                            SR                524289
  3
    192.168.12.2                                           1/1/c1/1:1000
192.168.17.2/32                            SR                524290
  3
    192.168.17.2                                           1/1/c2/1:1000
-------------------------------------------------------------------------------
Total Entries : 8
-------------------------------------------------------------------------------
===============================================================================
```

```
[/]
A:admin@PE-1# show router tunnel-table detail
```

```
===============================================================================
Tunnel Table (Router: Base)
===============================================================================
Destination      : 192.0.2.2/32
NextHop          : 192.168.12.2
Tunnel Flags     : entropy-label-capable
Age              : 00h01m06s
CBF Classes      : (Not Specified)
Owner            : isis (0)             Encap            : MPLS
Tunnel ID        : 524293               Preference       : 11
Tunnel Label     : 20002                Tunnel Metric    : 10
Tunnel MTU       : 8914                 Max Label Stack  : 1
-------------------------------------------------------------------------------
Destination      : 192.0.2.3/32
NextHop          : 192.168.12.2
Tunnel Flags     : entropy-label-capable
Age              : 00h01m06s
CBF Classes      : (Not Specified)
Owner            : isis (0)             Encap            : MPLS
Tunnel ID        : 524294               Preference       : 11
Tunnel Label     : 20003                Tunnel Metric    : 20
Tunnel MTU       : 8914                 Max Label Stack  : 1
-------------------------------------------------------------------------------
Destination      : 192.0.2.4/32 [L]
NextHop          : 192.168.12.2
Tunnel Flags     : has-lfa entropy-label-capable
Age              : 00h01m04s
CBF Classes      : (Not Specified)
Owner            : isis (0)             Encap            : MPLS
Tunnel ID        : 524295               Preference       : 11
Tunnel Label     : 20004                Tunnel Metric    : 30
Tunnel MTU       : 8914                 Max Label Stack  : 1
-------------------------------------------------------------------------------
Destination      : 192.0.2.5/32 [L]
NextHop          : 192.168.17.2
Tunnel Flags     : has-lfa entropy-label-capable
Age              : 00h01m04s
CBF Classes      : (Not Specified)
Owner            : isis (0)             Encap            : MPLS
Tunnel ID        : 524296               Preference       : 11
Tunnel Label     : 20005                Tunnel Metric    : 30
Tunnel MTU       : 8914                 Max Label Stack  : 1
-------------------------------------------------------------------------------
Destination      : 192.0.2.6/32
NextHop          : 192.168.17.2
Tunnel Flags     : entropy-label-capable
Age              : 00h01m06s
CBF Classes      : (Not Specified)
Owner            : isis (0)             Encap            : MPLS
Tunnel ID        : 524292               Preference       : 11
Tunnel Label     : 20006                Tunnel Metric    : 20
Tunnel MTU       : 8914                 Max Label Stack  : 1
-------------------------------------------------------------------------------
Destination      : 192.0.2.7/32
NextHop          : 192.168.17.2
Tunnel Flags     : entropy-label-capable
Age              : 00h01m06s
CBF Classes      : (Not Specified)
Owner            : isis (0)             Encap            : MPLS
Tunnel ID        : 524291               Preference       : 11
Tunnel Label     : 24007                Tunnel Metric    : 10
Tunnel MTU       : 8914                 Max Label Stack  : 1
-------------------------------------------------------------------------------
Destination      : 192.168.12.2/32
```

```
NextHop          : 192.168.12.2
Tunnel Flags     : is-adjacency-tunnel
Age              : 00h01m06s
CBF Classes      : (Not Specified)
Owner            : isis (0)              Encap            : MPLS
Tunnel ID        : 524289                Preference       : 11
Tunnel Label     : 3                     Tunnel Metric    : 0
Tunnel MTU       : 8914                  Max Label Stack  : 1
-------------------------------------------------------------------------------
Destination      : 192.168.17.2/32
NextHop          : 192.168.17.2
Tunnel Flags     : is-adjacency-tunnel
Age              : 00h01m06s
CBF Classes      : (Not Specified)
Owner            : isis (0)              Encap            : MPLS
Tunnel ID        : 524290                Preference       : 11
Tunnel Label     : 3                     Tunnel Metric    : 0
Tunnel MTU       : 8914                  Max Label Stack  : 1
-------------------------------------------------------------------------------
Number of tunnel-table entries          : 8
Number of tunnel-table entries with LFA : 2
===============================================================================
```

When a failure occurs on the primary SR path (only applicable for prefix PE-4/PE-5 and the link between PE-4 and PE-5), the traffic takes the LFA backup SR path to the destination using the same MPLS label value.

To extend the LFA FRR coverage, for example, to find an LFA protection for node PE-7, which is one of the VPRN service endpoints, RLFA can be enabled. RLFA creates a virtual LFA by using a repair tunnel to carry packets to a point in the network from where they will not be looped back to the source, but forwarded (SPF-based) toward the destination prefix.

The RLFA implementation uses the PQ algorithm. The node where RLFA is configured (PE-1 in this example) computes an extended P-space and a Q-space. The intersection of both spaces is called the PQ-node. This PQ node is the destination node of the repair tunnel using an SR shortest path tunnel. To compute both spaces, SPF is used.

In this example, IS-IS is used as the IGP, using a default metric value of 10 for all links. With the assumption that the link between PE-1 and PE-7 is broken, the calculation of both the extended P-space and the Q-space at PE-1 is as follows:

- extended P-space — An SPF computed from node PE-1 and rooted at PE-2. It is used to calculate the set of routers that are reachable without any path transiting the protected link between PE-1 and PE-7. The following nodes belong to the extended P-space: PE-2, PE-3, PE-4, and PE-5.
- Q-space — A reverse SPF computed from PE-1 and rooted from PE-7 (acting as destination proxy). It is used to calculate the set of routers that can reach PE-7 without transiting the protected link between PE-1 and PE-7. The nodes PE-4, PE-5, and PE-6 belong to the Q-space.

Possible PQ-nodes are PE-4 or PE-5, because they are in the intersection of both spaces.

RLFA is configured as follows:

```
# on PE-1:
configure {
    router "Base" {
        isis 0 {
            loopfree-alternate {
                remote-lfa
```

The following command shows the SR LFA coverage:

```
[/]
A:admin@PE-1# show router isis sr-lfa-coverage

===============================================================================
Rtr Base ISIS Instance 0 SR LFA Coverage
===============================================================================
MT-ID   SidType      Level Proto LFA       RLFA      TILFA     Coverage
-------------------------------------------------------------------------------
0       node-sid     L2    ipv4  2(33%)    4(66%)    0(0%)     6/6(100%)
0       adj-sid      L2    ipv4  0(0%)     2(100%)   0(0%)     2/2(100%)
===============================================================================
```

The nodes PE-2, PE-3, PE-6, and PE-7 now have RLFA protection, whereas PE-4 and PE-5 have LFA protection.

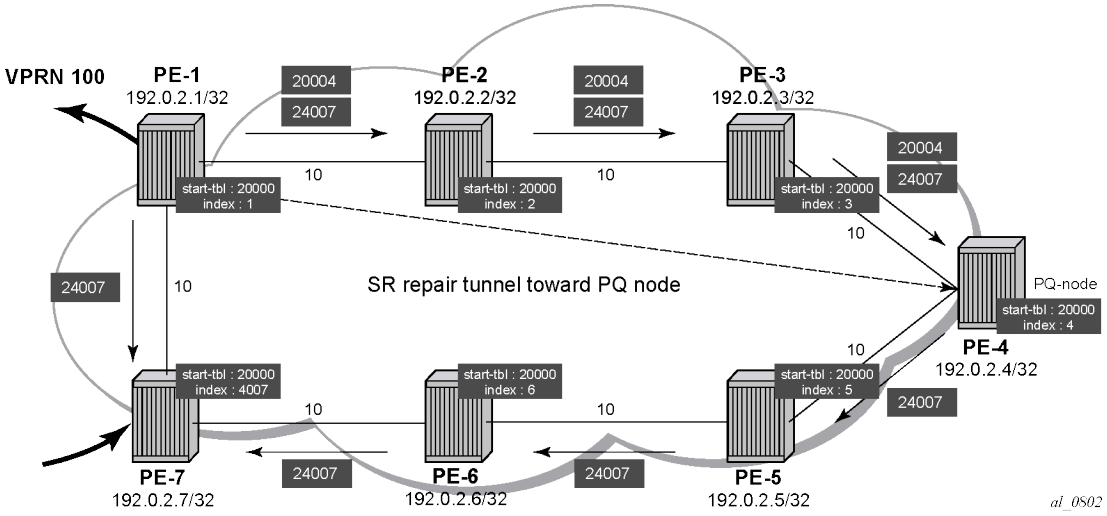
```
[/]
A:admin@PE-1# show router fp-tunnel-table 1

===============================================================================
IPv4 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                Protocol         Tunnel-ID
  Lbl/SID
    NextHop                                               Intf/Tunnel
  Lbl/SID (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
192.0.2.2/32                               SR-ISIS-0         524293
  20002
    192.168.12.2                                         1/1/c1/1:1000
  20002
  20005
    192.168.17.2(B)                                      1/1/c2/1:1000
192.0.2.3/32                               SR-ISIS-0         524294
  20003
    192.168.12.2                                         1/1/c1/1:1000
  20003
  20005
    192.168.17.2(B)                                      1/1/c2/1:1000
192.0.2.4/32                               SR-ISIS-0         524295
  20004
    192.168.12.2                                         1/1/c1/1:1000
  20004
    192.168.17.2(B)                                      1/1/c2/1:1000
192.0.2.5/32                               SR-ISIS-0         524296
  20005
    192.168.17.2                                         1/1/c2/1:1000
  20005
    192.168.12.2(B)                                      1/1/c1/1:1000
192.0.2.6/32                               SR-ISIS-0         524292
  20006
    192.168.17.2                                         1/1/c2/1:1000
  20006
  20004
    192.168.12.2(B)                                      1/1/c1/1:1000
192.0.2.7/32                               SR-ISIS-0         524291
  24007
```

```
    192.168.17.2                                        1/1/c2/1:1000
  24007
  20004
    192.168.12.2(B)                                     1/1/c1/1:1000
192.168.12.2/32                         SR                524289
  3
    192.168.12.2                                        1/1/c1/1:1000
  20002
  20005
    192.168.17.2(B)                                     1/1/c2/1:1000
192.168.17.2/32                         SR                524290
  3
    192.168.17.2                                        1/1/c2/1:1000
  24007
  20004
    192.168.12.2(B)                                     1/1/c1/1:1000
-------------------------------------------------------------------------------
Total Entries : 8
-------------------------------------------------------------------------------
===============================================================================
```

The main difference between normal prefix LFA and RLFA is that for RLFA a two-MPLS label stack is pushed by the head-end node (PE-1). The top label is the SR-label to reach the PQ node (for example, 20004 for PE-4) and the bottom label is the SR-label to reach the destination node (for example, 24007 for PE-7). The label stack inside the show command is ordered from bottom-most to top-most.

Figure 57: RLFA traffic path during protection illustrates the RLFA traffic path protecting the link between PE-1 and PE-7:

Figure 57: RLFA traffic path during protection



Inside the TTM, a tunnel-flag, *has-lfa*, is set for all destination nodes that have LFA protection available. The last two tunnels are adjacency tunnels and have in addition the flag *is-adjacency-tunnel*.

```
[/]
A:admin@PE-1# show router tunnel-table detail

===============================================================================
Tunnel Table (Router: Base)
===============================================================================
Destination      : 192.0.2.2/32 [L]
```

```
NextHop          : 192.168.12.2
Tunnel Flags     : has-lfa entropy-label-capable
Age              : 00h00m49s
CBF Classes      : (Not Specified)
Owner            : isis (0)             Encap            : MPLS
Tunnel ID        : 524293               Preference       : 11
Tunnel Label     : 20002                Tunnel Metric    : 10
Tunnel MTU       : 8910                 Max Label Stack  : 2
-------------------------------------------------------------------------------
Destination      : 192.0.2.3/32 [L]
NextHop          : 192.168.12.2
Tunnel Flags     : has-lfa entropy-label-capable
Age              : 00h00m49s
CBF Classes      : (Not Specified)
Owner            : isis (0)             Encap            : MPLS
Tunnel ID        : 524294               Preference       : 11
Tunnel Label     : 20003                Tunnel Metric    : 20
Tunnel MTU       : 8910                 Max Label Stack  : 2
-------------------------------------------------------------------------------
Destination      : 192.0.2.4/32 [L]
NextHop          : 192.168.12.2
Tunnel Flags     : has-lfa entropy-label-capable
Age              : 00h00m51s
CBF Classes      : (Not Specified)
Owner            : isis (0)             Encap            : MPLS
Tunnel ID        : 524295               Preference       : 11
Tunnel Label     : 20004                Tunnel Metric    : 30
Tunnel MTU       : 8910                 Max Label Stack  : 2
-------------------------------------------------------------------------------
Destination      : 192.0.2.5/32 [L]
NextHop          : 192.168.17.2
Tunnel Flags     : has-lfa entropy-label-capable
Age              : 00h00m51s
CBF Classes      : (Not Specified)
Owner            : isis (0)             Encap            : MPLS
Tunnel ID        : 524296               Preference       : 11
Tunnel Label     : 20005                Tunnel Metric    : 30
Tunnel MTU       : 8910                 Max Label Stack  : 2
-------------------------------------------------------------------------------
Destination      : 192.0.2.6/32 [L]
NextHop          : 192.168.17.2
Tunnel Flags     : has-lfa entropy-label-capable
Age              : 00h00m49s
CBF Classes      : (Not Specified)
Owner            : isis (0)             Encap            : MPLS
Tunnel ID        : 524292               Preference       : 11
Tunnel Label     : 20006                Tunnel Metric    : 20
Tunnel MTU       : 8910                 Max Label Stack  : 2
-------------------------------------------------------------------------------
Destination      : 192.0.2.7/32 [L]
NextHop          : 192.168.17.2
Tunnel Flags     : has-lfa entropy-label-capable
Age              : 00h00m49s
CBF Classes      : (Not Specified)
Owner            : isis (0)             Encap            : MPLS
Tunnel ID        : 524291               Preference       : 11
Tunnel Label     : 24007                Tunnel Metric    : 10
Tunnel MTU       : 8910                 Max Label Stack  : 2
-------------------------------------------------------------------------------
Destination      : 192.168.12.2/32 [L]
NextHop          : 192.168.12.2
Tunnel Flags     : has-lfa is-adjacency-tunnel
Age              : 00h00m49s
CBF Classes      : (Not Specified)
```

```
Owner            : isis (0)             Encap            : MPLS
Tunnel ID        : 524289               Preference       : 11
Tunnel Label     : 3                    Tunnel Metric    : 0
Tunnel MTU       : 8910                 Max Label Stack  : 2
-------------------------------------------------------------------------------
Destination      : 192.168.17.2/32 [L]
NextHop          : 192.168.17.2
Tunnel Flags     : has-lfa is-adjacency-tunnel
Age              : 00h00m49s
CBF Classes      : (Not Specified)
Owner            : isis (0)             Encap            : MPLS
Tunnel ID        : 524290               Preference       : 11
Tunnel Label     : 3                    Tunnel Metric    : 0
Tunnel MTU       : 8910                 Max Label Stack  : 2
-------------------------------------------------------------------------------
Number of tunnel-table entries          : 8
Number of tunnel-table entries with LFA : 8
===============================================================================
```

Verification of the loopback address configured within the VPRN service context on PE-7 (using loopback address 10.10.1.7/32) shows that an SR shortest path tunnel is used as the transport mechanism:

```
[/]
A:admin@PE-1# show router service-name "VPRN-100" route-table 10.10.1.7/32 extensive

===============================================================================
Route Table (Service: 100)
===============================================================================
Dest Prefix             : 10.10.1.7/32
  Protocol              : BGP_VPN
  Age                   : 00h02m07s
  Preference            : 170
  Indirect Next-Hop     : 192.0.2.7
    Label               : 524284
    VPN Next-Hop Index  : 10
    QoS                 : Priority=n/c, FC=n/c
    Source-Class        : 0
    Dest-Class          : 0
    ECMP-Weight         : N/A
    Resolving Next-Hop  : 192.0.2.7 (SR-ISIS tunnel:524291)
      Metric            : 10
      ECMP-Weight       : N/A
-------------------------------------------------------------------------------
No. of Destinations: 1
===============================================================================
```

In the following example, LFA/RLFA is no longer configured on the PE-1 node:

```
# on PE-1:
configure {
    router "Base" {
        isis 0 {
            delete loopfree-alternate
```

Example 2 - TTM preference with VPRN service

The following example is a variant on the previous example. The difference in this example is that, in addition to segment routing, LDP and RSVP-TE are also enabled between PE-1 and PE-7. A single RSVP LSP is configured originating at PE-1 and terminating at PE-7.

The objective of this example is to show the difference in protocol preference within TTM and how to influence the default behavior. This can be useful in case of migration scenarios from a non-SR environment toward a hybrid environment having LDP/RSVP and SR enabled.

```
# on PE-1:
configure {
    router "Base" {
        mpls {
            admin-state enable
            interface "int-PE-1-PE-7" {
            }
            path "dyn" {
                admin-state enable
            }
            lsp "LSP-PE-1-PE-7" {
                admin-state enable
                type p2p-rsvp
                to 192.0.2.7
                primary "dyn"
                }
            }
        }
        rsvp {
            admin-state enable
            interface "int-PE-1-PE-7" {
            }
        }
        ldp {
            interface-parameters {
                interface "int-PE-1-PE-7" {
                    ipv4 {
                    }
                }
            }
        }
```

```
# on PE-7:
configure {
    router "Base" {
        mpls {
            admin-state enable
            interface "int-PE-7-PE-1" {
            }
        }
        rsvp {
            admin-state enable
            interface "int-PE-7-PE-1" {
            }
        }
        ldp {
            interface-parameters {
                interface "int-PE-7-PE-1" {
                    ipv4 {
                    }
                }
            }
        }
```

By enabling LDP and RSVP between PE-1 and PE-7, the TTM on both nodes changed. With the VPRN service between PE-1 and PE-7 of example 1, only those two specific service endpoints are displayed:

```
[/]
A:admin@PE-1# show router tunnel-table 192.0.2.7

===============================================================================
IPv4 Tunnel Table (Router: Base)
===============================================================================
Destination           Owner     Encap TunnelId  Pref   Nexthop        Metric
   Color
-------------------------------------------------------------------------------
192.0.2.7/32          rsvp      MPLS  1         7      192.168.17.2   10
192.0.2.7/32          ldp       MPLS  65537     9      192.168.17.2   10
192.0.2.7/32          isis (0)  MPLS  524291    11     192.168.17.2   10
-------------------------------------------------------------------------------
Flags: B = BGP or MPLS backup hop available
       L = Loop-Free Alternate (LFA) hop available
       E = Inactive best-external BGP route
       k = RIB-API or Forwarding Policy backup hop
===============================================================================
```

```
[/]
A:admin@PE-7# show router tunnel-table 192.0.2.1

===============================================================================
IPv4 Tunnel Table (Router: Base)
===============================================================================
Destination           Owner     Encap TunnelId  Pref   Nexthop        Metric
   Color
-------------------------------------------------------------------------------
192.0.2.1/32          ldp       MPLS  65537     9      192.168.17.1   10
192.0.2.1/32          isis (0)  MPLS  524294    11     192.168.17.1   10
-------------------------------------------------------------------------------
Flags: B = BGP or MPLS backup hop available
       L = Loop-Free Alternate (LFA) hop available
       E = Inactive best-external BGP route
       k = RIB-API or Forwarding Policy backup hop
===============================================================================
```

On node PE-1, an RSVP LSP, an LDP LSP, and an SR shortest path tunnel (using IS-IS) are present. Because the VPRN service has **auto-bind-tunnel resolution any** enabled, the protocol type with the highest TTM preference (meaning the lowest absolute preference value in TTM) is taken; in this case, the RSVP LSP. This can be verified for the configured loopback address within the VPRN service context, as follows:

```
[/]
A:admin@PE-1# show router 100 route-table 10.10.1.7/32 extensive

===============================================================================
Route Table (Service: 100)
===============================================================================
Dest Prefix             : 10.10.1.7/32
  Protocol              : BGP_VPN
  Age                   : 00h01m37s
  Preference            : 170
  Indirect Next-Hop     : 192.0.2.7
    Label               : 524284
    VPN Next-Hop Index  : 10
    QoS                 : Priority=n/c, FC=n/c
    Source-Class        : 0
```

```
    Dest-Class          : 0
    ECMP-Weight         : N/A
    Resolving Next-Hop  : 192.0.2.7 (RSVP tunnel:1)
      Metric            : 10
      ECMP-Weight       : N/A
-------------------------------------------------------------------------------
No. of Destinations: 1
===============================================================================
```

On node PE-7, only an LDP LSP and an SR shortest path tunnel (using IS-IS) are present. Because the VPRN service has **auto-bind-tunnel resolution any** enabled, the protocol type with highest TTM preference (meaning the lowest absolute preference value in TTM) is taken; in this case, the LDP LSP. This can be verified for the configured loopback address within the VPRN service context, as follows:

```
[/]
A:admin@PE-7# show router 100 route-table 10.10.1.1/32 extensive

===============================================================================
Route Table (Service: 100)
===============================================================================
Dest Prefix             : 10.10.1.1/32
  Protocol              : BGP_VPN
  Age                   : 00h02m45s
  Preference            : 170
  Indirect Next-Hop     : 192.0.2.1
    Label               : 524284
    VPN Next-Hop Index  : 10
    QoS                 : Priority=n/c, FC=n/c
    Source-Class        : 0
    Dest-Class          : 0
    ECMP-Weight         : N/A
    Resolving Next-Hop  : 192.0.2.1 (LDP tunnel)
      Metric            : 10
      ECMP-Weight       : N/A
-------------------------------------------------------------------------------
No. of Destinations: 1
===============================================================================
```

Some configuration changes are possible to change this default behavior:

- It is possible to change the **auto-bind-tunnel resolution any** command into **auto-bind-tunnel resolution filter**. Because this is a service-specific parameter, the operator has the choice to only configure this on one specific service endpoint. From a migration point of view, a smooth and easy SR migration is possible, not affecting any other deployed services on this node.
- It is possible to change the SR tunnel-table protocol preference on a node. From a migration point of view, this affects all services initiating on this node.

Using the current example, PE-1 implements the auto-bind-tunnel change (option 1), while PE-7 implements the TTM preference change (option 2).

On PE-1, a **resolution-filter** CLI context within VPRN service 100 must be created. The example uses a **resolution-filter** context, which uses a filter to only allow SR shortest path tunnels (IS-IS based). The **auto-bind-tunnel resolution any** command is changed into **resolution filter**, as follows:

```
# on PE-1:
configure {
    service {
        vprn "VPRN-100" {
            bgp-ipvpn {
                mpls {
```

```
                    auto-bind-tunnel {
                        resolution filter
                        resolution-filter {
                            sr-isis true
                        }
                    }
```

As a result, the RSVP LSP is no longer used on PE-1. Instead, the SR shortest path tunnel is used for the traffic from PE-1 to PE-7:

```
[/]
A:admin@PE-1# show router 100 route-table 10.10.1.7/32 extensive

===============================================================================
Route Table (Service: 100)
===============================================================================
Dest Prefix             : 10.10.1.7/32
  Protocol              : BGP_VPN
  Age                   : 00h00m09s
  Preference            : 170
  Indirect Next-Hop     : 192.0.2.7
    Label               : 524284
    VPN Next-Hop Index  : 11
    QoS                 : Priority=n/c, FC=n/c
    Source-Class        : 0
    Dest-Class          : 0
    ECMP-Weight         : N/A
    Resolving Next-Hop  : 192.0.2.7 (SR-ISIS tunnel:524291)
      Metric            : 10
      ECMP-Weight       : N/A
-------------------------------------------------------------------------------
No. of Destinations: 1
===============================================================================
```

The VPRN service on node PE-7 is still using the LDP LSP as transport mechanism to reach node PE-1 at this point. Because the previous CLI change is only done within the VPRN service context 100 on PE-1, only the direction from PE-1 to PE-7 is affected.

Another way to influence the default TTM preference is shown as follows on the PE-7 node. Using the default behavior, the LDP LSP is used, because of the preference value of 9. If the SR tunnel table preference value is lowered to a value smaller than LDP, for instance 4, the SR shortest path tunnels originating on this node will always have preference compared to LDP LSP. On PE-7, the SR tunnel table preference is configured with a value of *4*, as follows:

```
# on PE-7:
configure {
    router "Base" {
        isis 0 {
            segment-routing {
                tunnel-table-pref 4
```

```
[/]
A:admin@PE-7# show router tunnel-table 192.0.2.1

===============================================================================
IPv4 Tunnel Table (Router: Base)
===============================================================================
Destination           Owner     Encap TunnelId  Pref   Nexthop        Metric
   Color
-------------------------------------------------------------------------------
```

```
192.0.2.1/32          isis (0)  MPLS  524291    4      192.168.17.1   10
192.0.2.1/32          ldp       MPLS  65537     9      192.168.17.1   10
-------------------------------------------------------------------------------
Flags: B = BGP or MPLS backup hop available
       L = Loop-Free Alternate (LFA) hop available
       E = Inactive best-external BGP route
       k = RIB-API or Forwarding Policy backup hop
===============================================================================
```

As a result, the LDP LSP is no longer used on PE-7 and the SR shortest path tunnel is the preferred transport tunnel:

```
[/]
A:admin@PE-7# show router service-name "VPRN-100" route-table 10.10.1.1/32 extensive

===============================================================================
Route Table (Service: 100)
===============================================================================
Dest Prefix             : 10.10.1.1/32
  Protocol              : BGP_VPN
  Age                   : 00h00m19s
  Preference            : 170
  Indirect Next-Hop     : 192.0.2.1
    Label               : 524284
    VPN Next-Hop Index  : 10
    QoS                 : Priority=n/c, FC=n/c
    Source-Class        : 0
    Dest-Class          : 0
    ECMP-Weight         : N/A
    Resolving Next-Hop  : 192.0.2.1 (SR-ISIS tunnel:524294)
      Metric            : 10
      ECMP-Weight       : N/A
-------------------------------------------------------------------------------
No. of Destinations: 1
===============================================================================
```

At this point, within the VPRN service, the SR shortest path tunnels are used bidirectionally between PE-1 and PE-7.

If, for example, an operator configures explicit SDP binding within the same VPRN service on both endpoints, the explicit SDPs will always have preference. In this example, manual SDPs are configured on nodes PE-1 and PE-7, both using LDP, as follows:

```
# on PE-1:
configure {
    service
        sdp 17 {
            admin-state enable
            delivery-type mpls
            ldp true
            far-end {
                ip-address 192.0.2.7
            }
        }
        vprn "VPRN-100" {
            spoke-sdp 17:100 {
            }
```

```
# on PE-7:
configure {
    service {
```

```
        sdp 71 {
            admin-state enable
            delivery-type mpls
            ldp true
            far-end {
                ip-address 192.0.2.1
            }
        }
        vprn "VPRN-100" {
            spoke-sdp 71:100 {
            }
```

As a result, SR shortest path tunnels are no longer used, but rather LDP-based SDPs are used instead:

```
[/]
A:admin@PE-1# show router 100 route-table 10.10.1.7/32 extensive

===============================================================================
Route Table (Service: 100)
===============================================================================
Dest Prefix             : 10.10.1.7/32
  Protocol              : BGP_VPN
  Age                   : 00h00m04s
  Preference            : 170
  Indirect Next-Hop     : 192.0.2.7
    Label               : 524284
    VPN Next-Hop Index  : 12
    QoS                 : Priority=n/c, FC=n/c
    Source-Class        : 0
    Dest-Class          : 0
    ECMP-Weight         : N/A
    Resolving Next-Hop  : 192.0.2.7 (SDP tunnel:17)
      Metric            : 0
      ECMP-Weight       : N/A
-------------------------------------------------------------------------------
No. of Destinations: 1
===============================================================================
```

```
[/]
A:admin@PE-7# show router 100 route-table 10.10.1.1/32 extensive

===============================================================================
Route Table (Service: 100)
===============================================================================
Dest Prefix             : 10.10.1.1/32
  Protocol              : BGP_VPN
  Age                   : 00h00m10s
  Preference            : 170
  Indirect Next-Hop     : 192.0.2.1
    Label               : 524284
    VPN Next-Hop Index  : 11
    QoS                 : Priority=n/c, FC=n/c
    Source-Class        : 0
    Dest-Class          : 0
    ECMP-Weight         : N/A
    Resolving Next-Hop  : 192.0.2.1 (SDP tunnel:71)
      Metric            : 0
      ECMP-Weight       : N/A
-------------------------------------------------------------------------------
No. of Destinations: 1
===============================================================================
```

Conclusion

Segment Routing is a technique using extensions of the existing link state protocols, and using existing MPLS or IPv6 infrastructure as the data plane. It is a source routing technique similar to RSVP-TE, but without the need to run an extra signaling protocol. SR also avoids other scaling restrictions of associated RSVP-TE, such as midpoint state. SR is simple to control and operate because the intelligence and state are part of the packet, not held by the network. Other benefits are that SR can be introduced in an incremental way using different migration scenarios to assure a smooth transition.

SR-TE LSP Path Computation Using Local CSPF

This chapter describes the SR-TE LSP path computation using local CSPF.

Topics in this chapter include:

Applicability

This chapter was initially written for SR OS Release 19.10.R1, but the MD-CLI in the current edition corresponds to SR OS Release 25.10.R2. Local CSPF can be used in IPv4 SR-TE LSP primary and secondary path computation in SR OS Release 19.7.R1, and later.

Overview

Segment Routing with Traffic Engineering Label Switched Paths (SR-TE LSPs) can be computed using:

- hop-to-label (IP-to-label) translation (default; **path-computation-method** not specified)
- Path Computation Element (PCE) path computation (**path-computation-method pce**)
- local Constrained Shortest Path First (CSPF) (**path-computation-method local-cspf**)

Hop-to-label path computation

SR-TE LSP path computation uses hop-to-label translation as the default computation method. The PCC interrogates the TE database, and translates any hop configured in the applied path statement to a Node SID (N-SID) or Adjacency SID (A-SID), to produce a list of segment IDs. Strict hops are mapped to adjacency SIDs; loose hops are mapped to node SIDs. The destination address in the LSP configuration implies a final loose hop.

PCE path computation

SR-TE LSP path computation can also be performed using an external PCE controller. In this case, the PCC maintains a Path Computation Element Protocol (PCEP) session with the PCE and the path computation is done as follows:

- the PCC sends a PCReq requesting a path
- the PCE replies with a PCReply including a path (if available). This path contains a segment list.

- Optionally, the PCC sends a path status report (PCRpt) to the PCE. However, the PCC may also delegate the control of the path to the PCE.

PCE path computation is supported for SR-TE LSPs, but not for SR-TE LSP templates. You cannot have PCE path computation for SR-TE LSPs that use LSP templates **one-hop-p2p-srte** or **mesh-p2p-srte** auto-LSPs. PCE path computation is not further treated in this chapter.

Local CSPF path computation

SR-TE LSP path computation using local CSPF can be used in single-area OSPFv2 or single-level IS-IS IGP instances. More complex LSP path computations, or when the network is expanded into multiple IGP areas or instances, require an external PCE.

One of the major changes to the SR-TE algorithm from RSVP-TE CSPF is that SR-TE does not require each router to be TE enabled: the links do not have to be TE links. Provided that the routers at each end of the link are SR enabled, local CSPF computes an end-to-end path.

Full CSPF path computation on the head-end router (PCC) results in a full explicit path to the destination. The PCC calculates an end-to-end path and the following applies:

- The computed path is a full explicit TE path.
- Each link is represented by an adjacency SID or adjacency set SID.
- CSPF returns a label stack list of adjacency SIDs or adjacency set SIDs.

Like RSVP-TE LSPs, an SR-TE LSP can be resignaled when a timer expires or when an operator issues a tools command.

Paths computed by local CSPF contain an adjacency SID for each link in the path and the stack may contain numerous labels. If the **max-sr-labels** value may be exceeded or the maximum segment depth of a downstream router may be less than the calculated LSP label stack size, the label stack can be reduced. The label reduction capability can replace a series of adjacency SIDs with a node SID. For loose-hop path computation, node SIDs can be used or a combination of node and adjacency SIDs.

Local CSPF is supported on both primary and secondary standby paths of an SR-TE LSP. Local CSPF path calculation can be used for RSVP-TE LSP as well as for SR-TE LSP templates.

Local CSPF path computation and SR protected interfaces

When SR is enabled and IGP adjacency is established over a link, the router advertises an adjacency SID in the adjacency SID sub-TLV. When Loop-Free Alternate (LFA), Remote LFA (RLFA), or Topology-Independent LFA (TI-LFA) is enabled, protected adjacencies have the backup flag (B-flag) set in the adjacency SID sub-TLV. Each adjacency is available for SID protection when LFA, RLFA, or TI-LFA is enabled. It is possible to remove the SID protection on a specific link (**sid-protection false**).

Adjacency sets are specified in an adjacency-set sub-TLV as a single object. Adjacency sets never have the B-flag set and are always unprotected. However, each individual link in the adjacency set is protected. For more information about adjacency sets, see the Parallel Adjacency Sets in Segment Routing chapter.

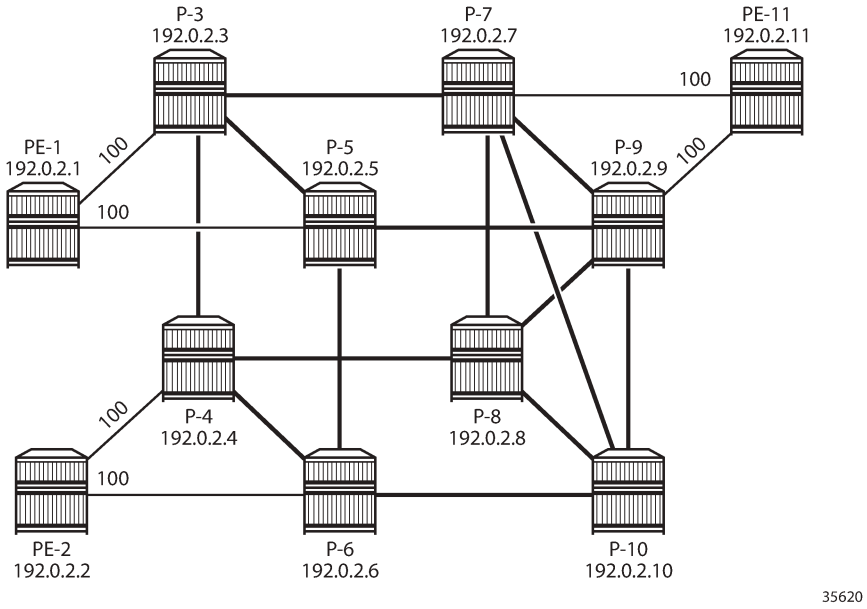
Local CSPF path calculation can set up a path that:

- only includes protected adjacencies (**local-sr-protection mandatory**)
- only includes unprotected adjacencies (**local-sr-protection none**)
- can include both protected and unprotected adjacencies (**local-sr-protection preferred** (default))

Configuration

Figure 58: Example topology shows the example topology.

Figure 58: Example topology



The initial configuration on each of the nodes includes:

- Cards, MDAs, ports
- Router interfaces
- IS-IS enabled on all router interfaces (alternatively, OSPFv2 can be used as IGP)
  - The interfaces in the core (between P-3, P-4, P-5, P-6, P-7, P-8, P-9, and P-10) have metric 10.
  - The access interfaces to and from PE-1, PE-2, and PE-11 have metric 100.
  - TE is enabled on the head-end routers.
- MPLS is enabled on the head-end routers.

For an in-depth description of the configuration of SR-ISIS, see the Segment Routing with IS-IS Control Plane chapter. On PE-2, the IS-IS configuration is as follows:

```
# on PE-2:
configure {
    router "Base" {
        mpls-labels {
            sr-labels {
                start 32000
                end 32999
            }
        }
        isis 0 {
            admin-state enable
            advertise-router-capability area
```

```
            level-capability 2                      # single-level IS-IS instance
            traffic-engineering true                # TE enabled in headend
            area-address [49.0001]
            segment-routing {
                admin-state enable
                prefix-sid-range {
                    global
                }
            }
            interface "int-PE-2-P-4" {
                interface-type point-to-point
                level 2 {
                    metric 100                      # metric 100 on access interfaces
                }
            }
            interface "int-PE-2-P-6" {
                interface-type point-to-point
                level 2 {
                    metric 100                      # metric 100 on access interfaces
                }
            }
            interface "system" {
                ipv4-node-sid {
                    index 2
                }
            }
            level 2 {
                wide-metrics-only true
            }
        }
```

With this configuration, the node SID on PE-2 is 32000 + index 2 = 32002. The configuration is similar on the other nodes.

On PE-2, the following SR-TE LSPs are configured toward PE-11:

- SR-TE LSP with empty path and:
  - hop-to-label path computation
  - local CSPF path computation without label stack reduction
  - local CSPF path computation with label stack reduction
- SR-TE LSP with path with two strict hops—P-4 and P-3—and:
  - hop-to-label path computation
  - local CSPF path computation without label stack reduction
  - local CSPF path computation with label stack reduction
- SR-TE LSP with path with two loose hops—P-3 and P-9—and:
  - hop-to-label path computation
  - local CSPF path computation without label stack reduction
  - local CSPF path computation with label stack reduction

SR-TE LSPs using empty path

The configuration of SR-TE LSPs is described in chapter Segment Routing – Traffic Engineered Tunnels chapter. On PE-2, the following SR-TE LSPs toward PE-11 are configured with an empty path. The path computation method is hop-to-label for the first SR-TE LSP and local CSPF for the second SR-TE LSP.

```
# on PE-2:
configure {
    router "Base" {
        mpls {
            admin-state enable
            path "empty-path" {
                admin-state enable
            }
            lsp "LSP-PE-2-PE-11_empty-path_hop-to-label" {
                admin-state enable
                type p2p-sr-te
                to 192.0.2.11
                max-sr-labels {
                    additional-frr-labels 2
                }
                primary "empty-path" {
                }
            }
            lsp "LSP-PE-2-PE-11_empty-path_localCSPF" {
                admin-state enable
                type p2p-sr-te
                to 192.0.2.11
                path-computation-method local-cspf
                max-sr-labels {
                    additional-frr-labels 2
                }
                primary "empty-path" {
                }
            }
```

With hop-to-label path computation, the destination 192.0.2.11 is an implied loose hop that is mapped to the node SID 32011 of the destination PE-11, as follows:

```
[/]
A:admin@PE-2# show router mpls sr-te-lsp "LSP-PE-2-PE-11_empty-path_hop-to-label" path detail

===============================================================================
MPLS SR-TE LSP LSP-PE-2-PE-11_empty-path_hop-to-label
Path  (Detail)
===============================================================================
Legend :
    S      - Strict                     L      - Loose
    A-SID  - Adjacency SID              N-SID  - Node SID
    +      - Inherited
===============================================================================
-------------------------------------------------------------------------------
LSP SR-TE LSP-PE-2-PE-11_empty-path_hop-to-label
Path  empty-path
-------------------------------------------------------------------------------
LSP Name    : LSP-PE-2-PE-11_empty-path_hop-to-label
Path LSP ID      : 32256
From             : 192.0.2.2
To               : 192.0.2.11
Admin State      : Up                       Oper State        : Up
```

```
Path Name   : empty-path
Path Type         : Primary
Path Admin        : Up                        Path Oper         : Up
Path Up Time      : 0d 00:31:57               Path Down Time    : 0d 00:00:00
Retry Limit       : 0                         Retry Timer       : 30 sec
Retry Attempt     : 0                         Next Retry In     : 0 sec

PathCompMethod    : none                      OperPathCompMethod: none
MetricType        : igp                       Oper MetricType   : igp
LocalSrProt       : preferred                 Oper LocalSrProt  : N/A
LabelStackRed     : Disabled                  Oper LabelStackRed: N/A

Bandwidth         : No Reservation            Oper Bandwidth    : 0 Mbps
Hop Limit         : 255                       Oper HopLimit     : 255
Setup Priority    : 7                         Oper SetupPriority: 7
Hold Priority     : 0                         Oper HoldPriority : 0
DelayMetricLimit : No limit                   OperDelayMetricLim: N/A
Inter-area        : N/A

PCE Updt ID       : 0                         PCE Updt State    : None
PCE Upd Fail Code: noError

PCE Report        : Disabled+                 Oper PCE Report   : Disabled
PCE Control       : Disabled                  Oper PCE Control  : Disabled

Include Groups    :                           Oper IncludeGroups:
None                                              None
Exclude Groups    :                           Oper ExcludeGroups:
None                                              None
Last Resignal     : n/a

IGP/TE/Del Metric: 16777215                   Oper Metric       : 16777215
Oper MTU          : 8970                      Path Trans        : 1
Degraded          : False
Failure Code      : noError
Failure Node      : n/a
Explicit Hops     :
    No Hops Specified
Actual Hops       :
    192.0.2.11(192.0.2.11)(N-SID)                 Record Label        : 32011

BFD Configuration and State
Template          : None                      Ping Interval     : N/A
Enable            : False                     State             : notApplicable
ReturnPathLabel   : None
BFD Source Addr   : None
WaitForUpTimer    : 4 sec                     OperWaitForUpTimer: 0 sec
WaitForUpTmLeft   : 0
StartFail Rsn     : N/A

===============================================================================
```

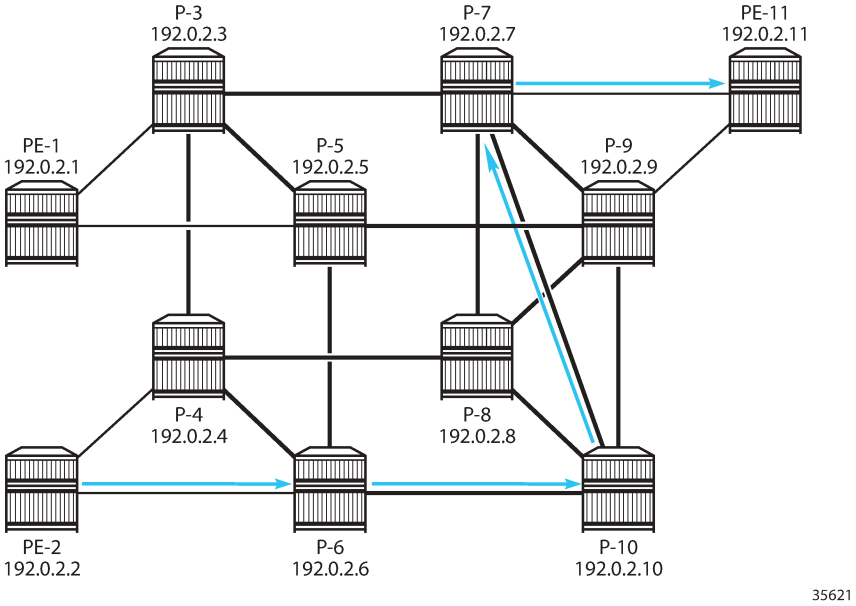
With local CSPF path computation, the SR-TE path contains contiguous strict hops with A-SIDs. Several ECMP paths are available and, in this case, the path goes from PE-2 via P-6, P-10, and P-7 to the destination PE-11, as follows:

```
[/]
A:admin@PE-2# show router mpls sr-te-lsp "LSP-PE-2-PE-11_empty-path_localCSPF" path detail

===============================================================================
MPLS SR-TE LSP LSP-PE-2-PE-11_empty-path_localCSPF
Path  (Detail)
===============================================================================
```

```
Legend :
    S      - Strict                      L      - Loose
    A-SID  - Adjacency SID               N-SID  - Node SID
    +      - Inherited
===============================================================================
-------------------------------------------------------------------------------
LSP SR-TE LSP-PE-2-PE-11_empty-path_localCSPF
Path  empty-path
-------------------------------------------------------------------------------
LSP Name    : LSP-PE-2-PE-11_empty-path_localCSPF
Path LSP ID      : 50176
From             : 192.0.2.2
To               : 192.0.2.11
Admin State      : Up                        Oper State        : Up
Path Name   : empty-path
Path Type        : Primary
Path Admin       : Up                        Path Oper         : Up
Path Up Time     : 0d 00:31:57               Path Down Time    : 0d 00:00:00
Retry Limit      : 0                         Retry Timer       : 30 sec
Retry Attempt    : 0                         Next Retry In     : 0 sec

PathCompMethod   : local-cspf                OperPathCompMethod: local-cspf
MetricType       : igp                       Oper MetricType   : igp
LocalSrProt      : preferred                 Oper LocalSrProt  : preferred
LabelStackRed    : Disabled                  Oper LabelStackRed: Disabled

---snip---

IGP/TE/Del Metric: 220                       Oper Metric       : 220
Oper MTU         : 8958                      Path Trans        : 1
Degraded         : False
Failure Code     : noError
Failure Node     : n/a
Explicit Hops    :
    No Hops Specified
Actual Hops      :
    192.168.26.2(192.0.2.6)(A-SID)               Record Label        : 524286
 -> 192.168.106.2(192.0.2.10)(A-SID)             Record Label        : 524284
 -> 192.168.107.1(192.0.2.7)(A-SID)              Record Label        : 524287
 -> 192.168.117.2(192.0.2.11)(A-SID)             Record Label        : 524283

---snip---
```

The path goes from PE-2 via P-6, P-10, and P-7 to the destination PE-11, as shown in Figure 59: Empty path from PE-2 to PE-11.

Figure 59: Empty path from PE-2 to PE-11



By enabling label stack reduction, the label stack of Adjacency SIDs (A-SIDs) can be reduced to a smaller number of Node SIDs (N-SIDs), or a combination of N-SIDs and A-SIDs, while still honoring TE constraints.

If the user enables label stack reduction for an SR-TE LSP, a second CSPF phase is applied, attempting to reduce the label stack that resulted from the fully explicit path with adjacency SIDs and adjacency sets SIDs computed in CSPF phase 1. The procedure of the label stack reduction algorithm is described in the 7750 SR and 7950 XRS Segment Routing and PCE User Guide.

The following SR TE LSP is configured with **label-stack-reduction true**:

```
# on PE-2:
configure {
    router "Base" {
        mpls {
            lsp "LSP-PE-2-PE-11_empty-path_localCSPF" {
                admin-state enable
                type p2p-sr-te
                to 192.0.2.11
                path-computation-method local-cspf
                label-stack-reduction true
                max-sr-labels {
                    additional-frr-labels 2
                }
                primary "empty-path" {
                }
            }
```

In this example, the label stack is reduced to the N-SID of the destination PE-11, as follows:

```
[/]
A:admin@PE-2# show router mpls sr-te-lsp "LSP-PE-2-PE-11_empty-path_localCSPF_red" path detail

---snip---
```

```
-------------------------------------------------------------------------------
LSP SR-TE LSP-PE-2-PE-11_empty-path_localCSPF_red
Path  empty-path
-------------------------------------------------------------------------------
---snip---

PathCompMethod   : local-cspf              OperPathCompMethod: local-cspf
MetricType       : igp                     Oper MetricType   : igp
LocalSrProt      : preferred               Oper LocalSrProt  : preferred
LabelStackRed    : Enabled                 Oper LabelStackRed: Enabled

---snip---

IGP/TE/Del Metric: 220                     Oper Metric       : 220
Oper MTU         : 8970                    Path Trans        : 1
Degraded         : False
Failure Code     : noError
Failure Node     : n/a
Explicit Hops    :
    No Hops Specified
Actual Hops      :
    192.0.2.11(192.0.2.11)(N-SID)                Record Label        : 32011

---snip---
```

SR-TE LSPs using path with strict hops

In the following example, the SR-TE LSP path includes strict hops (P-4 and P-3)—that must be contiguous hops from the headend router—and an implicit loose hop to the destination 192.0.2.11. The SR-TE LSPs on PE-2 are configured as follows:

```
# on PE-2:
configure {
    router "Base" {
        mpls {
            path "path-via-P-4-P-3_S" {
                admin-state enable
                hop 10 {
                    ip-address 192.0.2.4
                    type strict
                }
                hop 20 {
                    ip-address 192.0.2.3
                    type strict
                }
            }
            lsp "LSP-PE-2-PE-11_w-strict-hops_hop-to-label" {
                admin-state enable
                type p2p-sr-te
                to 192.0.2.11
                max-sr-labels {
                    additional-frr-labels 2
                }
                primary "path-via-P-4-P-3_S" {
                }
            }
            lsp "LSP-PE-2-PE-11_w-strict-hops_localCSPF" {
                admin-state enable
                type p2p-sr-te
                to 192.0.2.11
                path-computation-method local-cspf
```

```
                max-sr-labels {
                    additional-frr-labels 2
                }
                primary "path-via-P-4-P-3_S" {
                }
            }
```

With hop-to-label path computation, strict hops are translated into adjacency SIDs, whereas loose hops are translated into node SIDs. In this example, the path has an A-SID to P-4 and an A-SID to P-3 followed by an N-SID to the destination PE-11, as follows:

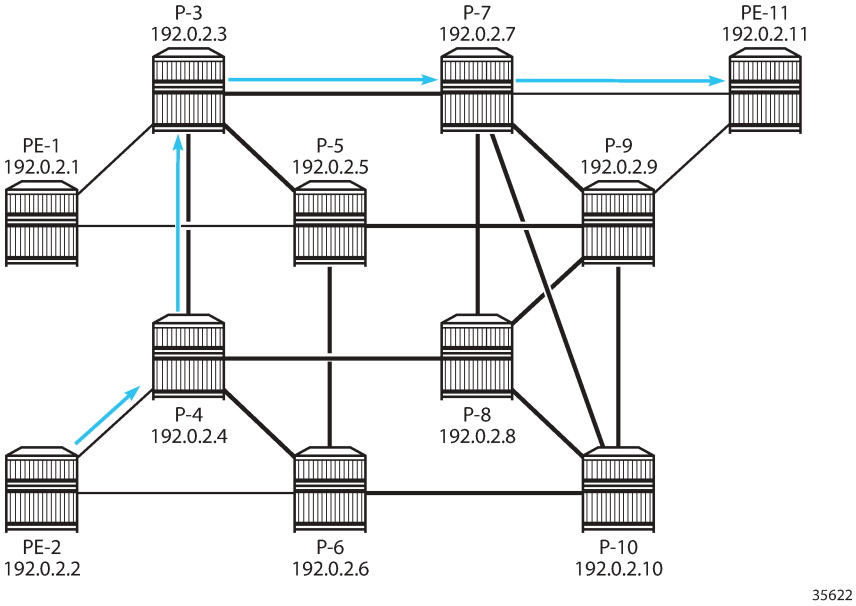
```
[/]
A:admin@PE-2# show router mpls sr-te-lsp "LSP-PE-2-PE-11_w-strict-hops_hop-to-label" path
 detail

===============================================================================
MPLS SR-TE LSP LSP-PE-2-PE-11_w-strict-hops_hop-to-label
Path  (Detail)
===============================================================================
Legend :
    S      - Strict                      L      - Loose
    A-SID  - Adjacency SID               N-SID  - Node SID
    +      - Inherited
===============================================================================
-------------------------------------------------------------------------------
LSP SR-TE LSP-PE-2-PE-11_w-strict-hops_hop-to-label
Path  path-via-P-4-P-3_S
-------------------------------------------------------------------------------
LSP Name    : LSP-PE-2-PE-11_w-strict-hops_hop-to-label
Path LSP ID      : 49664
From             : 192.0.2.2
To               : 192.0.2.11
Admin State      : Up                        Oper State        : Up
Path Name   : path-via-P-4-P-3_S
Path Type        : Primary
Path Admin       : Up                        Path Oper         : Up
Path Up Time     : 0d 00:32:30               Path Down Time    : 0d 00:00:00
Retry Limit      : 0                         Retry Timer       : 30 sec
Retry Attempt    : 0                         Next Retry In     : 0 sec

PathCompMethod   : none                      OperPathCompMethod: none
MetricType       : igp                       Oper MetricType   : igp
LocalSrProt      : preferred                 Oper LocalSrProt  : N/A
LabelStackRed    : Disabled                  Oper LabelStackRed: N/A

---snip---

IGP/TE/Del Metric: 16777215                  Oper Metric       : 16777215
Oper MTU         : 8962                      Path Trans        : 1
Degraded         : False
Failure Code     : noError
Failure Node     : n/a
Explicit Hops    :
                  192.0.2.4(S)
               -> 192.0.2.3(S)
Actual Hops      :
    192.168.24.2(192.0.2.4)(A-SID)                Record Label        : 524287
 -> 192.168.34.1(192.0.2.3)(A-SID)                Record Label        : 524286
 -> 192.0.2.11(192.0.2.11)(N-SID)                 Record Label        : 32011

---snip---
```

With local CSPF path computation, the path is a sequence of A-SIDs, as follows:

```
[/]
A:admin@PE-2# show router mpls sr-te-lsp "LSP-PE-2-PE-11_w-strict-hops_localCSPF" path detail

===============================================================================
MPLS SR-TE LSP LSP-PE-2-PE-11_w-strict-hops_localCSPF
Path  (Detail)
===============================================================================
Legend :
    S      - Strict                      L      - Loose
    A-SID  - Adjacency SID               N-SID  - Node SID
    +      - Inherited
===============================================================================
-------------------------------------------------------------------------------
LSP SR-TE LSP-PE-2-PE-11_w-strict-hops_localCSPF
Path  path-via-P-4-P-3_S
-------------------------------------------------------------------------------
LSP Name    : LSP-PE-2-PE-11_w-strict-hops_localCSPF
Path LSP ID      : 51200
From             : 192.0.2.2
To               : 192.0.2.11
Admin State      : Up                        Oper State        : Up
Path Name   : path-via-P-4-P-3_S
Path Type        : Primary
Path Admin       : Up                        Path Oper         : Up
Path Up Time     : 0d 00:32:30               Path Down Time    : 0d 00:00:00
Retry Limit      : 0                         Retry Timer       : 30 sec
Retry Attempt    : 0                         Next Retry In     : 0 sec

PathCompMethod   : local-cspf                OperPathCompMethod: local-cspf
MetricType       : igp                       Oper MetricType   : igp
LocalSrProt      : preferred                 Oper LocalSrProt  : preferred
LabelStackRed    : Disabled                  Oper LabelStackRed: Disabled

---snip---

IGP/TE/Del Metric: 220                       Oper Metric       : 220
Oper MTU         : 8958                      Path Trans        : 1
Degraded         : False
Failure Code     : noError
Failure Node     : n/a
Explicit Hops    :
                    192.0.2.4(S)
                 -> 192.0.2.3(S)
Actual Hops      :
    192.168.24.2(192.0.2.4)(A-SID)                 Record Label        : 524287
 -> 192.168.34.1(192.0.2.3)(A-SID)                 Record Label        : 524286
 -> 192.168.37.2(192.0.2.7)(A-SID)                 Record Label        : 524284
 -> 192.168.117.2(192.0.2.11)(A-SID)               Record Label        : 524283

---snip---
```

The path from PE-2 to PE-11 must go via P-4 and P-3. The loose hop from P-3 to the destination PE-11 is translated into an A-SID to P-7 followed by an A-SID to PE-11, as shown in Figure 60: Path from PE-2 to PE-11 via strict hops P-4 and P-3.

Figure 60: Path from PE-2 to PE-11 via strict hops P-4 and P-3



With label stack reduction, the configuration of the SR-TE LSP is as follows:

```
# on PE-2:
configure {
    router "Base" {
        mpls {
            lsp "LSP-PE-2-PE-11_w-strict-hops_localCSPF_red" {
                admin-state enable
                type p2p-sr-te
                to 192.0.2.11
                path-computation-method local-cspf
                label-stack-reduction true
                max-sr-labels {
                    additional-frr-labels 2
                }
                primary "path-via-P-4-P-3_S" {
                }
            }
```

Label stack reduction reduces the label stack to one or more node SIDs in segments, with each segment delimited by configured path hops. The computed path to the node SID must satisfy any required path constraints. The calculated path from PE-2 to PE-11 via the strict hops P-4 and P-3 shows a series of N-SIDs, as follows:

```
[/]
A:admin@PE-2# show router mpls sr-te-lsp "LSP-PE-2-PE-11_w-strict-hops_localCSPF_red" path
 detail

===============================================================================
MPLS SR-TE LSP LSP-PE-2-PE-11_w-strict-hops_localCSPF_red
Path  (Detail)
===============================================================================
Legend :
    S      - Strict                      L      - Loose
```

```
    A-SID  - Adjacency SID              N-SID  - Node SID
    +      - Inherited
===============================================================================
---snip---

PathCompMethod   : local-cspf             OperPathCompMethod: local-cspf
MetricType       : igp                    Oper MetricType   : igp
LocalSrProt      : preferred              Oper LocalSrProt  : preferred
LabelStackRed    : Enabled                Oper LabelStackRed: Enabled

---snip---

IGP/TE/Del Metric: 220                    Oper Metric       : 220
Oper MTU         : 8962                   Path Trans        : 1
Degraded         : False
Failure Code     : noError
Failure Node     : n/a
Explicit Hops    :
                   192.0.2.4(S)
                -> 192.0.2.3(S)
Actual Hops      :
    192.0.2.4(192.0.2.4)(N-SID)                 Record Label        : 32004
 -> 192.0.2.3(192.0.2.3)(N-SID)                 Record Label        : 32003
 -> 192.0.2.11(192.0.2.11)(N-SID)               Record Label        : 32011

---snip---
```

SR-TE LSPs using path with loose hops

The following SR-TE LSPs on PE-2 toward PE-11 use a path with loose hops P-3 and P-9:

```
# on PE-2:
configure {
    router "Base" {
        mpls {
            path "path-via-P-3-P-9_L" {
                admin-state enable
                hop 10 {
                    ip-address 192.0.2.3
                    type loose
                }
                hop 20 {
                    ip-address 192.0.2.9
                    type loose
                }
            }
            lsp "LSP-PE-2-PE-11_w-loose-hops_hop-to-label" {
                admin-state enable
                type p2p-sr-te
                to 192.0.2.11
                max-sr-labels {
                    additional-frr-labels 2
                }
                primary "path-via-P-3-P-9_L" {
                }
            }
            lsp "LSP-PE-2-PE-11_w-loose-hops_localCSPF" {
                admin-state enable
                type p2p-sr-te
                to 192.0.2.11
                path-computation-method local-cspf
```

```
                max-sr-labels {
                    additional-frr-labels 2
                }
                primary "path-via-P-3-P-9_L" {
                }
            }
```

With hop-to-label path calculation, loose hops are translated into N-SIDs. In this example, the actual hops are the N-SIDs of P-3, P-9, and PE-11, as follows:

```
[/]
A:admin@PE-2# show router mpls sr-te-lsp "LSP-PE-2-PE-11_w-loose-hops_hop-to-label" path detail

===============================================================================
MPLS SR-TE LSP LSP-PE-2-PE-11_w-loose-hops_hop-to-label
Path  (Detail)
===============================================================================
Legend :
    S      - Strict                     L      - Loose
    A-SID  - Adjacency SID              N-SID  - Node SID
    +      - Inherited
===============================================================================
-------------------------------------------------------------------------------
LSP SR-TE LSP-PE-2-PE-11_w-loose-hops_hop-to-label
Path  path-via-P-3-P-9_L
-------------------------------------------------------------------------------
---snip---

PathCompMethod   : none                   OperPathCompMethod: none
MetricType       : igp                    Oper MetricType   : igp
LocalSrProt      : preferred              Oper LocalSrProt  : N/A
LabelStackRed    : Disabled               Oper LabelStackRed: N/A

---snip---

IGP/TE/Del Metric: 16777215               Oper Metric       : 16777215
Oper MTU         : 8962                   Path Trans        : 1
Degraded         : False
Failure Code     : noError
Failure Node     : n/a
Explicit Hops    :
                   192.0.2.3(L)
                -> 192.0.2.9(L)
Actual Hops      :
    192.0.2.3(192.0.2.3)(N-SID)                  Record Label        : 32003
 -> 192.0.2.9(192.0.2.9)(N-SID)                  Record Label        : 32009
 -> 192.0.2.11(192.0.2.11)(N-SID)                Record Label        : 32011

---snip---
```

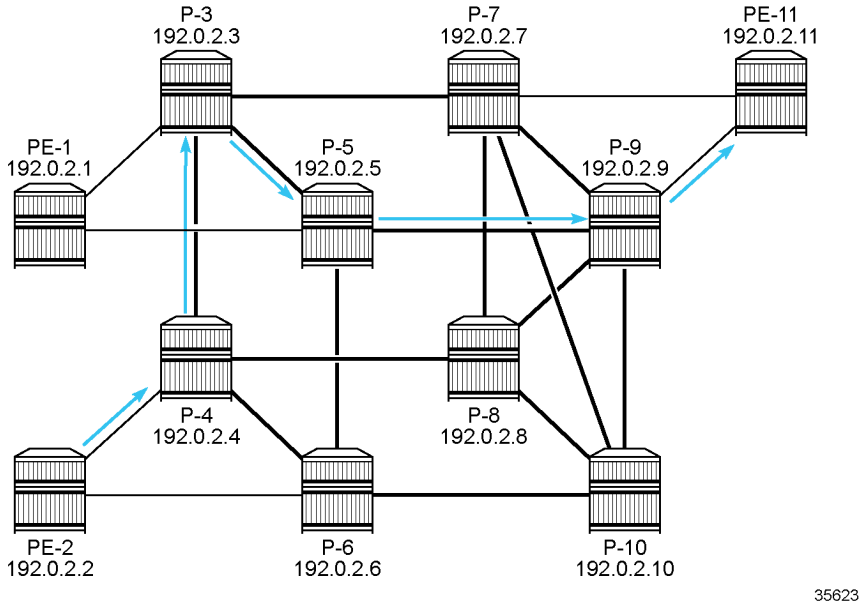
With local CSPF path calculation, the actual hops are the A-SIDs toward P-4, P-3, P-5, P-9, and PE-11, as follows:

```
[/]
A:admin@PE-2# show router mpls sr-te-lsp "LSP-PE-2-PE-11_w-loose-hops_localCSPF" path detail

===============================================================================
MPLS SR-TE LSP LSP-PE-2-PE-11_w-loose-hops_localCSPF
Path  (Detail)
===============================================================================
Legend :
    S      - Strict                     L      - Loose
```

```
    A-SID  - Adjacency SID               N-SID  - Node SID
    +      - Inherited
===============================================================================
-------------------------------------------------------------------------------
LSP SR-TE LSP-PE-2-PE-11_w-loose-hops_localCSPF
Path  path-via-P-3-P-9_L
-------------------------------------------------------------------------------
LSP Name    : LSP-PE-2-PE-11_w-loose-hops_localCSPF
Path LSP ID      : 9728
From             : 192.0.2.2
To               : 192.0.2.11
Admin State      : Up                      Oper State        : Up
Path Name   : path-via-P-3-P-9_L
Path Type        : Primary
Path Admin       : Up                      Path Oper         : Up
Path Up Time     : 0d 00:32:57             Path Down Time    : 0d 00:00:00
Retry Limit      : 0                       Retry Timer       : 30 sec
Retry Attempt    : 0                       Next Retry In     : 0 sec

PathCompMethod   : local-cspf              OperPathCompMethod: local-cspf
MetricType       : igp                     Oper MetricType   : igp
LocalSrProt      : preferred               Oper LocalSrProt  : preferred
LabelStackRed    : Disabled                Oper LabelStackRed: Disabled

---snip---

IGP/TE/Del Metric: 230                     Oper Metric       : 230
Oper MTU         : 8954                    Path Trans        : 1
Degraded         : False
Failure Code     : noError
Failure Node     : n/a
Explicit Hops    :
                  192.0.2.3(L)
               -> 192.0.2.9(L)
Actual Hops      :
    192.168.24.2(192.0.2.4)(A-SID)               Record Label        : 524287
 -> 192.168.34.1(192.0.2.3)(A-SID)               Record Label        : 524286
 -> 192.168.35.2(192.0.2.5)(A-SID)               Record Label        : 524285
 -> 192.168.59.2(192.0.2.9)(A-SID)               Record Label        : 524284
 -> 192.168.119.2(192.0.2.11)(A-SID)             Record Label        : 524283

---snip---
```

Figure 61: Path from PE-2 to PE-11 via loose hops P-3 and P-9 shows the path from PE-2 to PE-11 via loose hops P-3 and P-9.

Figure 61: Path from PE-2 to PE-11 via loose hops P-3 and P-9



Label stack reduction is configured as follows:

```
# on PE-2:
configure {
    router "Base" {
        mpls {
            lsp "LSP-PE-2-PE-11_w-loose-hops_localCSPF_red" {
                admin-state enable
                type p2p-sr-te
                to 192.0.2.11
                path-computation-method local-cspf
                label-stack-reduction true
                max-sr-labels {
                    additional-frr-labels 2
                }
                primary "path-via-P-3-P-9_L" {
                }
            }
```

With label stack reduction, the actual hops in the path are the following:

```
[/]
A:admin@PE-2# show router mpls sr-te-lsp "LSP-PE-2-PE-11_w-loose-hops_localCSPF_red" path
 detail

===============================================================================
MPLS SR-TE LSP LSP-PE-2-PE-11_w-loose-hops_localCSPF_red
Path  (Detail)
===============================================================================
Legend :
    S      - Strict                      L      - Loose
    A-SID  - Adjacency SID               N-SID  - Node SID
    +      - Inherited
===============================================================================
```

```
-------------------------------------------------------------------------------
LSP SR-TE LSP-PE-2-PE-11_w-loose-hops_localCSPF_red
Path  path-via-P-3-P-9_L
-------------------------------------------------------------------------------
---snip---

PathCompMethod   : local-cspf             OperPathCompMethod: local-cspf
MetricType       : igp                    Oper MetricType   : igp
LocalSrProt      : preferred              Oper LocalSrProt  : preferred
LabelStackRed    : Enabled                Oper LabelStackRed: Enabled

---snip---

IGP/TE/Del Metric: 230                    Oper Metric       : 230
Oper MTU         : 8962                   Path Trans        : 1
Degraded         : False
Failure Code     : noError
Failure Node     : n/a
Explicit Hops    :
                    192.0.2.3(L)
                -> 192.0.2.9(L)
Actual Hops      :
    192.0.2.3(192.0.2.3)(N-SID)                 Record Label        : 32003
 -> 192.0.2.9(192.0.2.9)(N-SID)                 Record Label        : 32009
 -> 192.0.2.11(192.0.2.11)(N-SID)               Record Label        : 32011

---snip---
```

Tunnel tables

The following command on PE-2 lists the SR-TE tunnels.

```
[/]
A:admin@PE-2# show router tunnel-table protocol sr-te

===============================================================================
IPv4 Tunnel Table (Router: Base)
===============================================================================
Destination           Owner     Encap TunnelId  Pref   Nexthop        Metric
   Color
-------------------------------------------------------------------------------
192.0.2.11/32         sr-te     MPLS  655363    8      192.168.26.2   220
192.0.2.11/32         sr-te     MPLS  655364    8      192.0.2.11     220
192.0.2.11/32         sr-te     MPLS  655369    8      192.168.24.2   220
192.0.2.11/32         sr-te     MPLS  655370    8      192.0.2.4      220
192.0.2.11/32         sr-te     MPLS  655366    8      192.168.24.2   230
192.0.2.11/32         sr-te     MPLS  655367    8      192.0.2.3      230
192.0.2.11/32         sr-te     MPLS  655362    8      192.0.2.11     16777215
192.0.2.11/32         sr-te     MPLS  655365    8      192.0.2.3      16777215
192.0.2.11/32         sr-te     MPLS  655368    8      192.168.24.2   16777215
-------------------------------------------------------------------------------
Flags: B = BGP or MPLS backup hop available
       L = Loop-Free Alternate (LFA) hop available
       E = Inactive best-external BGP route
       k = RIB-API or Forwarding Policy backup hop
===============================================================================
```

The SR-TE LSPs correspond to the tunnel IDs as follows:

- tunnel ID 655362: LSP with empty path and hop-to-label computation

- tunnel ID 655363: LSP with empty path and local CSPF computation without label stack reduction
- tunnel ID 655364: LSP with empty path and local CSPF computation with label stack reduction
- tunnel ID 655365: LSP with loose hops and hop-to-label computation
- tunnel ID 655366: LSP with loose hops and local CSPF computation without label stack reduction
- tunnel ID 655367: LSP with loose hops and local CSPF computation with label stack reduction
- tunnel ID 655368: LSP with strict hops and hop-to-label computation
- tunnel ID 655369: LSP with strict hops and local CSPF computation without label stack reduction
- tunnel ID 655370: LSP with strict hops and local CSPF computation with label stack reduction

By default, SR-TE tunnels have preference 8. LSPs 655362, 655365, and 655368 use hop-to-label path computation and have metric 16777215. For hop-to-label path computation, only the strict hops are translated into adjacency SIDs, which is the case for the LSP with tunnel ID 655368.

The three SR-TE LSP tunnels with metric 220 or 230 and next-hop 192.168.24.2 or 192.168.26.2 use local CSPF path computation without label stack reduction, while the three SR-TE tunnels with metric 220 or 230 and next hop equal to a system IP address 192.0.2.x use local CSPF path computation with label stack reduction. For all SR-TE LSPs with next-hop 192.168.24.2 or 192.168.26.2, the first hop is mapped to an adjacency SID. All paths computed with local CSPF without label stack reduction only have adjacency SIDs or adjacency set SIDs. When label stack reduction is configured, the next hops are slightly different when the first hop in the reduced label stack is mapped to a node SID.

For the hop-to-label computed paths, the value of the metric is set to 16777215 (infinity – 1), because CSPF is not used and the head-end router is unaware of the full topology between head- and tail-end router. For the paths computed with local CSPF, the IGP metrics are added; for example, for the LSP tunnel with tunnel ID 655363 (empty path with local CSPF): 100 (PE-2 to P-6) + 10 (P-6 to P-10) + 10 (P-10 to P-7) + 100 (P-7 to PE-11) = 220.

The following command on PE-2 shows the FP tunnel table for the SR-TE LSP tunnels:

```
[/]
A:admin@PE-2# show router fp-tunnel-table 1 protocol sr-te

===============================================================================
IPv4 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                Protocol         Tunnel-ID
  Lbl/SID
    NextHop                                               Intf/Tunnel
  Lbl/SID (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
192.0.2.11/32                              SR-TE             655362
  3
    192.0.2.11                                           SR
192.0.2.11/32                              SR-TE             655363
  524283
  524287
  524284
    192.168.26.2                                         SR
192.0.2.11/32                              SR-TE             655364
  3
    192.0.2.11                                           SR
```

```
192.0.2.11/32                                SR-TE             655365
  32011
  32009
    192.0.2.3                                                SR
192.0.2.11/32                                SR-TE             655366
  524283
  524284
  524285
  524286
    192.168.24.2                                             SR
192.0.2.11/32                                SR-TE             655367
  32011
  32009
    192.0.2.3                                                SR
192.0.2.11/32                                SR-TE             655368
  32011
  524286
    192.168.24.2                                             SR
192.0.2.11/32                                SR-TE             655369
  524283
  524284
  524286
    192.168.24.2                                             SR
192.0.2.11/32                                SR-TE             655370
  32011
  32003
    192.0.2.4                                                SR
-------------------------------------------------------------------------------
Total Entries : 9
-------------------------------------------------------------------------------
===============================================================================
```

The SR-TE LSP tunnels with tunnel IDs 655362, 655365, and 655368 have hop-to-label path computation and the other six SR-TE LSPs have local CSPF path computation. For hop-to-label path computation, A-SIDs are used for strict hops and N-SIDs are used for loose hops. For local CSPF path computation without label stack reduction, only A-SIDs and adjacency set SIDs are used. For local CSPF path computation with label stack reduction, N-SIDs replace sequences of A-SIDs.

Resignaling an SR-TE LSP

Point-to-point SR-TE LSPs have a resignal timer to match that of RSVP. It must be set to allow manual and automatic resignaling for optimization of SR-TE LSPs. The following command can be used to set a resignal timer in minutes for all originating SR-TE LSPs:

```
[ex:/configure router "Base" mpls sr-te-resignal]
A:admin@PE-2# resignal-timer ?

 resignal-timer <number>
 <number>  - <30..10080>  - minutes

    Resignal timer for SR-TE LSPs
```

The following tools command can be launched for manually-triggered re-optimization of LSPs:

```
[/]
A:admin@PE-2# tools perform router mpls resignal ?

 resignal p2mp-lsp <string> p2mp-instance <string>
 resignal p2mp-delay <number>
```

```
 resignal lsp <string> path <string>
 resignal delay <number>
 resignal sr-te-lsp <string> path <string>
 resignal sr-te-delay <number>

 delay                 - <number> - <0..30>
 lsp                   - string '<1..64 characters>'
 p2mp-delay            - <number> - <0..60>
 p2mp-lsp              - string '<1..64 characters>'
 sr-te-delay           - <number> - <0..30>
 sr-te-lsp             - string '<1..64 characters>'
```

A manual re-optimization for a specific path in a specific SR-TE LSP can be forced, as follows:

```
[/]
A:admin@PE-2# tools perform router mpls resignal sr-te-lsp "LSP-PE-2-PE-11_w-loose-hops_
localCSPF" path "path-via-P-3-P-9_L"
```

The **sr-te-delay** parameter overrides the global resignal timer value for all SR-TE LSPs. When this timer expires, the procedures of the timer-based resignal are applied to all SR-TE LSPs and the SR-TE resignal time is then reset to its configured value in the MPLS configuration.

The following command forces a re-optimization of the SR-TE LSPs after an **sr-te-delay** of 3 minutes, but this CLI delay only takes effect when the **sr-te-resignal-timer** is configured in the **mpls** context. If not, the following error is raised:

```
[/]
A:admin@PE-2# tools perform router mpls resignal sr-te-delay 3
WARNING: CLI #2006: Warning while processing command - WARNING: CLI Delay will not be in
 effect, configure resignal-timer under config>router>mpls>sr-te-resignal
```

The SR-TE resignal timer is configured in the **mpls** context with a value of 60 minutes, as follows:

```
# on PE-2:
configure {
    router "Base" {
        mpls {
            sr-te-resignal {
                resignal-timer 60
            }
```

With the SR-TE resignal timer configured, the tools command can be launched to override this SR-TE resignal timer to a value of 3 minutes, as follows:

```
[/]
A:admin@PE-2# tools perform router mpls resignal sr-te-delay 3
```

Local CSPF and SR protected adjacencies

The following command enables TI-LFA with link protection on all nodes:

```
# on all nodes:
configure {
    router "Base" {
        isis 0 {
            loopfree-alternate {
                ti-lfa {
```

```
            }
        }
```

As a result of this, each adjacency is available for SID protection. For example, on P-4, all adjacency SID sub-TLVs have the B-flag set, as follows:

```
[/]
A:admin@P-4# show router isis database P-4 detail level 2 | match "Adj-SID" pre-lines 1
    Nbr IP    : 192.168.24.1
    Adj-SID: Flags:v4BVL Weight:0 Label:524287
    Nbr IP    : 192.168.34.1
    Adj-SID: Flags:v4BVL Weight:0 Label:524286
    Nbr IP    : 192.168.46.2
    Adj-SID: Flags:v4BVL Weight:0 Label:524285
    Nbr IP    : 192.168.48.2
    Adj-SID: Flags:v4BVL Weight:0 Label:524284
                E = Entropy Label Capability (ELC) Flag
Adj-SID Flags    : v4/v6 = IPv4 or IPv6 Address-Family
                  B = Backup Flag
                  V = Adj-SID carries a value
```

The following command removes SID protection from the interface toward P-8:

```
# on P-4:
configure {
    router "Base" {
        isis 0 {
            interface "int-P-4-P-8" {
                sid-protection false
            }
```

The adjacency SID sub-TLV for this link does not have the B-flag set, as follows:

```
[/]
A:admin@P-4# show router isis database P-4 detail level 2

===============================================================================
Rtr Base ISIS Instance 0 Database (detail)
===============================================================================

Displaying Level 2 database
-------------------------------------------------------------------------------
LSP ID    : P-4.00-00                                 Level     : L2
Sequence  : 0x19                  Checksum  : 0x559a   Lifetime  : 1188
Version   : 1                     Pkt Type  : 20       Pkt Ver   : 1
Attributes: L1L2                  Max Area  : 3        Alloc Len : 1492
SYS ID    : 1920.0000.2004        SysID Len : 6        Used Len  : 645

TLVs :

---snip---
  TE IS Nbrs   :
    Nbr   : P-8.00
    Default Metric  : 10
    Sub TLV Len     : 19
    IF Addr   : 192.168.48.1
    Nbr IP    : 192.168.48.2
    Adj-SID: Flags:v4VL Weight:0 Label:524284
  ---snip---

Level (2) LSP Count : 1
-------------------------------------------------------------------------------
```

```
---snip---
Adj-SID Flags    : v4/v6 = IPv4 or IPv6 Address-Family
                   B = Backup Flag
                   V = Adj-SID carries a value
                   L = value/index has local significance
                   S = Set of Adjacencies
                   P = Persistently allocated
---snip---
```

Local CSPF computes, by default, an end-to-end path by selecting protected adjacencies, which have the B-flag set, as previously described. If no such path is available, the local CSPF may select an unprotected adjacency with the assumption that all other path constraints are met.

The following **tools** command calculates the path from P-4 to P-8 without establishing it. With the **preferred** option, protected adjacencies are preferred over unprotected adjacency when both exist for a TE link. In this example, the shortest path contains the direct link to P-8, which does not have a backup:

```
[/]
A:admin@P-4# tools perform router mpls sr-te-cspf to 192.0.2.8 path-computation-method local-
cspf local-sr-protection preferred
Req CSPF TE path
    From: this node To: 192.0.2.8
CSPF TE Path
 To: 192.0.2.8
 [1] Source Add 192.0.2.4    Cost 10
   Hop  1 -> Label 524284 NH 192.168.48.1 --> 192.168.48.2 (192.0.2.8) Cost 10 Color 0x0
```

The following tools command calculates the path from P-4 to P-8 but all adjacencies must be protected, so the unprotected direct link to P-8 is excluded and the path goes via P-6 and P-10 to P-8:

```
[/]
A:admin@P-4# tools perform router mpls sr-te-cspf to 192.0.2.8 path-computation-method local-
cspf local-sr-protection mandatory
Req CSPF TE path
    From: this node To: 192.0.2.8
CSPF TE Path
 To: 192.0.2.8
 [1] Source Add 192.0.2.4    Cost 30
   Hop  1 -> Label 524285 NH 192.168.46.1 --> 192.168.46.2 (192.0.2.6) Cost 10 Color 0x0
   Hop  2 -> Label 524284 NH 192.168.106.1 --> 192.168.106.2 (192.0.2.10) Cost 10 Color 0x0
   Hop  3 -> Label 524286 NH 192.168.108.2 --> 192.168.108.1 (192.0.2.8) Cost 10 Color 0x0
```

The following **tools** command calculates the path from P-4 to P-8 with the restriction that each link in the path is unprotected. The shortest path is the unprotected link to P-8:

```
[/]
A:admin@P-4# tools perform router mpls sr-te-cspf to 192.0.2.8 path-computation-method local-
cspf local-sr-protection none
Req CSPF TE path
    From: this node To: 192.0.2.8
CSPF TE Path
 To: 192.0.2.8
 [1] Source Add 192.0.2.4    Cost 10
   Hop  1 -> Label 524284 NH 192.168.48.1 --> 192.168.48.2 (192.0.2.8) Cost 10 Color 0x0
```

On PE-2, the following SR-TE LSPs with local CSPF path computation using a loose path are configured toward PE-11:

- an SR-TE LSP with local SR protection preferred (= default setting)

- an SR-TE LSP with mandatory local SR protection, where all adjacencies must have a backup
- an SR-TE LSP without local SR protection (none), where all adjacencies are unprotected

```
# on PE-2:
configure {
    router "Base" {
        mpls {
            lsp "LSP-PE-2-PE-11_empty-path_localCSPF_protection-preferred" {
                admin-state enable
                type p2p-sr-te
                to 192.0.2.11
                path-computation-method local-cspf
                local-sr-protection preferred       # default
                max-sr-labels {
                    additional-frr-labels 2
                }
                primary "empty-path" {
                }
            }
            lsp "LSP-PE-2-PE-11_empty-path_localCSPF_protection-mandatory" {
                admin-state enable
                type p2p-sr-te
                to 192.0.2.11
                path-computation-method local-cspf
                local-sr-protection mandatory     # all links in E2E path protected
                max-sr-labels {
                    additional-frr-labels 2
                }
                primary "empty-path" {
                }
            }
            lsp "LSP-PE-2-PE-11_empty-path_localCSPF_protection-none" {
                admin-state enable
                type p2p-sr-te
                to 192.0.2.11
                path-computation-method local-cspf
                local-sr-protection none        # all links in E2E path unprotected
                max-sr-labels {
                    additional-frr-labels 2
                }
                primary "empty-path" {
                }
            }
```

For test purposes, the metric on the unprotected interface between P-4 and P-8 is lowered to 5, so the shortest path to PE-11 includes the unprotected interface when allowed:

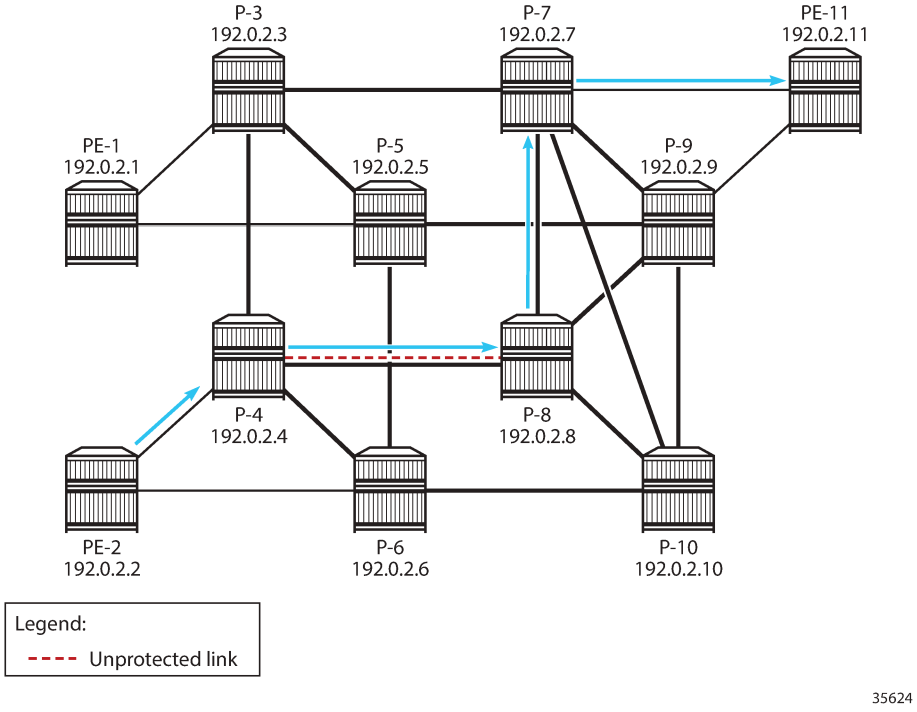
```
# on P-4:
configure {
     router "Base" {
        isis 0 {
            interface "int-P-4-P-8" {
                interface-type point-to-point
                sid-protection false
                level 2 {
                    metric 5
                }
            }
```

For SR-TE LSP "LSP-PE-2-PE-11_empty-path_localCSPF_protection-preferred", a path can be established from PE-2 via P-4, P-8, and P-7 to PE-11, but for the same metric, P-7 can be replaced by P-9. The direct link between P-4 and P-8 is unprotected, while all other adjacencies in the path have a backup.

```
[/]
A:admin@PE-2# show router mpls sr-te-lsp "LSP-PE-2-PE-11_empty-path_localCSPF_protection-
preferred" path detail

===============================================================================
MPLS SR-TE LSP LSP-PE-2-PE-11_empty-path_localCSPF_protection-preferred
Path  (Detail)
===============================================================================
Legend :
    S      - Strict                      L      - Loose
    A-SID  - Adjacency SID               N-SID  - Node SID
    +      - Inherited
===============================================================================
-------------------------------------------------------------------------------
LSP SR-TE LSP-PE-2-PE-11_empty-path_localCSPF_protection-preferred
Path  empty-path
-------------------------------------------------------------------------------
LSP Name    : LSP-PE-2-PE-11_empty-path_localCSPF_protection-preferred
Path LSP ID      : 58880
From             : 192.0.2.2
To               : 192.0.2.11
Admin State      : Up                        Oper State        : Up
Path Name   : empty-path
Path Type        : Primary
Path Admin       : Up                        Path Oper         : Up
Path Up Time     : 0d 00:00:16               Path Down Time    : 0d 00:00:00
Retry Limit      : 0                         Retry Timer       : 30 sec
Retry Attempt    : 0                         Next Retry In     : 0 sec

---snip---

IGP/TE/Del Metric: 215                       Oper Metric       : 215
Oper MTU         : 8958                      Path Trans        : 1
Degraded         : False
Failure Code     : noError
Failure Node     : n/a
Explicit Hops    :
    No Hops Specified
Actual Hops      :
    192.168.24.2(192.0.2.4)(A-SID)                Record Label        : 524287
 -> 192.168.48.2(192.0.2.8)(A-SID)                Record Label        : 524284    # unprotected
 adjacency
 -> 192.168.78.1(192.0.2.7)(A-SID)                Record Label        : 524286
 -> 192.168.117.2(192.0.2.11)(A-SID)              Record Label        : 524283
---snip---
```

Figure 62: Path from PE-2 to PE-11 including unprotected link shows the established path from PE-2 to PE-11, which uses protected and unprotected links.

Figure 62: Path from PE-2 to PE-11 including unprotected link



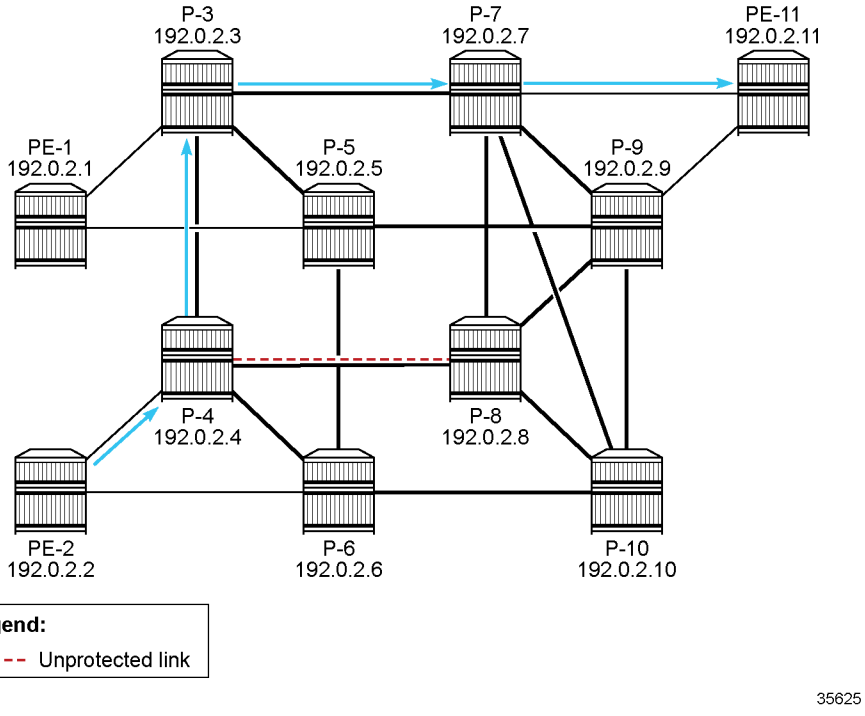
For SR-TE LSP "LSP-PE-2-PE-11_empty-path_localCSPF_protection-mandatory", the path must exclude the unprotected link from P-4 to P-8. The path goes from PE-2 via P-4, P-3, and P-7 to PE-11, but for the same metric, other paths are possible.

```
[/]
A:admin@PE-2# show router mpls sr-te-lsp "LSP-PE-2-PE-11_empty-path_localCSPF_protection-
mandatory" path detail

===============================================================================
MPLS SR-TE LSP LSP-PE-2-PE-11_empty-path_localCSPF_protection-mandatory
Path  (Detail)
===============================================================================
Legend :
    S      - Strict                      L      - Loose
    A-SID  - Adjacency SID               N-SID  - Node SID
    +      - Inherited
===============================================================================
-------------------------------------------------------------------------------
LSP SR-TE LSP-PE-2-PE-11_empty-path_localCSPF_protection-mandatory
Path  empty-path
-------------------------------------------------------------------------------
LSP Name    : LSP-PE-2-PE-11_empty-path_localCSPF_protection-mandatory
Path LSP ID      : 8704
From             : 192.0.2.2
To               : 192.0.2.11
Admin State      : Up                        Oper State        : Up
Path Name   : empty-path
Path Type        : Primary
Path Admin       : Up                        Path Oper         : Up
Path Up Time     : 0d 00:00:48               Path Down Time    : 0d 00:00:00
Retry Limit      : 0                         Retry Timer       : 30 sec
```

```
Retry Attempt    : 0                        Next Retry In     : 0 sec

PathCompMethod   : local-cspf               OperPathCompMethod: local-cspf
MetricType       : igp                      Oper MetricType   : igp
LocalSrProt      : mandatory                Oper LocalSrProt  : mandatory
LabelStackRed    : Disabled                 Oper LabelStackRed: N/A

---snip---

IGP/TE/Del Metric: 220                      Oper Metric       : 220
Oper MTU         : 8958                     Path Trans        : 1
Degraded         : False
Failure Code     : noError
Failure Node     : n/a
Explicit Hops    :
    No Hops Specified
Actual Hops      :
    192.168.24.2(192.0.2.4)(A-SID)               Record Label        : 524287
 -> 192.168.34.1(192.0.2.3)(A-SID)               Record Label        : 524286
 -> 192.168.37.2(192.0.2.7)(A-SID)               Record Label        : 524284
 -> 192.168.117.2(192.0.2.11)(A-SID)             Record Label        : 524283

---snip---
```

Figure 63: Path from PE-2 to PE-11 including only protected links shows the loose path that excludes the unprotected link between P-4 and P-8.

Figure 63: Path from PE-2 to PE-11 including only protected links



For SR-TE LSP "LSP-PE-2-PE-11_empty-path_localCSPF_protection-none", the path should only use unprotected links, but only the adjacency between P-4 and P-8 is unprotected, so CSPF

cannot find an end-to-end path. The path and the LSP remain operationally down with failure code noCspfRouteToDestination, as follows:

```
[/]
A:admin@PE-2# show router mpls sr-te-lsp "LSP-PE-2-PE-11_empty-path_localCSPF_protection-none"
 path detail

===============================================================================
MPLS SR-TE LSP LSP-PE-2-PE-11_empty-path_localCSPF_protection-none
Path  (Detail)
===============================================================================
Legend :
    S      - Strict                     L      - Loose
    A-SID  - Adjacency SID              N-SID  - Node SID
    +      - Inherited
===============================================================================
-------------------------------------------------------------------------------
LSP SR-TE LSP-PE-2-PE-11_empty-path_localCSPF_protection-none
Path  empty-path
-------------------------------------------------------------------------------
LSP Name    : LSP-PE-2-PE-11_empty-path_localCSPF_protection-none
Path LSP ID      : 26624
From             : 192.0.2.2
To               : 192.0.2.11
Admin State      : Up                       Oper State        : Down
Path Name   : empty-path
Path Type        : Primary
Path Admin       : Up                       Path Oper         : Down
Path Up Time     : 0d 00:00:00              Path Down Time    : 0d 00:01:52
Retry Limit      : 0                        Retry Timer       : 30 sec
Retry Attempt    : 5                        Next Retry In     : 14 sec

PathCompMethod   : local-cspf               OperPathCompMethod: N/A
MetricType       : igp                      Oper MetricType   : N/A
LocalSrProt      : none                     Oper LocalSrProt  : N/A
LabelStackRed    : Disabled                 Oper LabelStackRed: N/A

---snip---

IGP/TE/Del Metric: N/A                      Oper Metric       : N/A
Oper MTU         : N/A                      Path Trans        : 0
Degraded         : False
Failure Code     : noCspfRouteToDestination
Failure Node     : 192.0.2.2
Explicit Hops    :
    No Hops Specified
Actual Hops      :
    No Hops Specified

---snip---
```

Conclusion

Within a single-level IS-IS network or a single-area OSPF network, SR-TE LSP path calculation using local CSPF on the head-end router results in an end-to-end path using IPv4 adjacencies. The local CSPF path computation method can also be used for RSVP-TE LSPs.

SRv6 Encapsulation in the Base Routing Instance

This chapter provides information about SRv6 encapsulation in the base routing instance.

Topics in this chapter include:

Applicability

The information and configuration in this chapter are based on SR OS Release 22.2.R1. Segment Routing over IPv6 (SRv6) is supported on FP4-based equipment in SR OS Release 21.5.R2 and later.

Overview

SRv6 encapsulation in the base routing instance allows the transport of native IPv4 and IPv6 data across an SRv6-enabled network. To this end, native IPv4 and IPv6 data is sent to an ingress SRv6 router, where it is encapsulated and forwarded via an SRv6 tunnel. The SRv6 tunnel transports the encapsulated data across the SRv6-enabled network to an egress SRv6 router, where it is decapsulated and forwarded further as native IPv4 and IPv6 data. SRv6-tunneled data is encapsulated using an IPv6 header, where the destination address is a unique SRv6 segment identifier (SID), and is processed and forwarded in the IPv6 data plane.

An SRv6 SID is a preconfigured 128-bit routable IPv6 prefix address that is encoded in three parts: a locator, a function, and an argument. The locator is a summary IPv6 prefix for a set of SRv6 SIDs instantiated on an SRv6-capable router. It is used to route the data within the IPv6 transport network. Each participating SRv6-capable router needs its unique locator, based on a common block that all participating SRv6-capable routers share in the IPv6 address space. The function is an opaque identifier that indicates the local behavior at the endpoint of an SRv6 segment. The focus in this topic is on the SRv6 End.DT4 and the SRv6 End.DT6 functions, performing a prefix lookup in the global IPv4 route table (End.DT4) or in the global IPv6 route table (End.DT6). The argument is not used in SR OS 22.2.R1 and is set to all zeros.

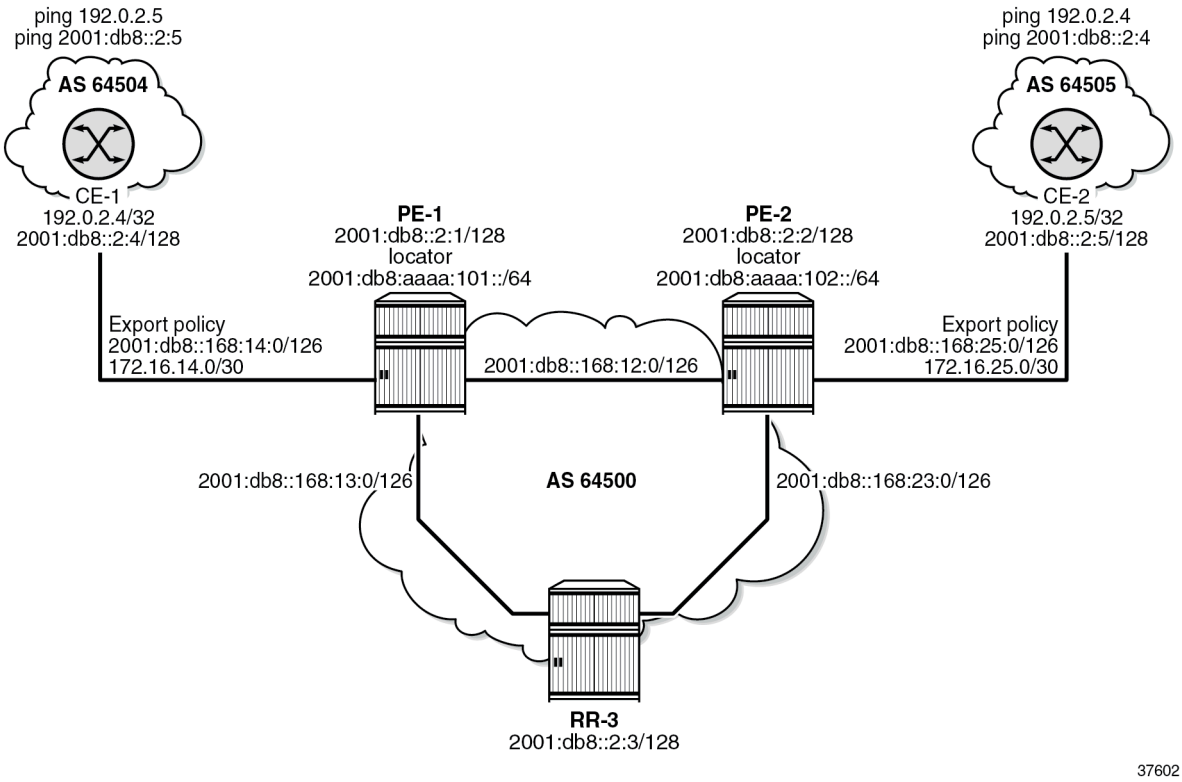
The local router installs its locator prefix in its IPv6 route table and Forwarding Information Base (FIB), and advertises its locator prefix in IS-IS with the SRv6 locator sub-TLV. Each remote router populates its route table and FIB with the received locator prefixes, including the tunneled next hop to the originating router.

SRv6 data transport requires additional processing at both the ingress and egress data planes. This processing relies on Forwarding Path Extension (FPE), as described in the Segment Routing over IPv6 chapter.

Configuration

Figure 64: Example topology shows the example topology with five routers. Two routers (CE-1 and CE-2) simulate an IPv6-enabled network. They are connected to an SRv6-enabled network, comprising of PE-1 and PE-2, and a route reflector (RR) RR-3 in the control plane. The SRv6-enabled network has only IPv6 addresses and interfaces.

Figure 64: Example topology



For the transport of native IPv4 and IPv6 data from CE-1 to CE-2, PE-1 acts as the SRv6 ingress PE node, while PE-2 acts as the SRv6 egress PE node. For the transport of native IPv4 and IPv6 data from CE-2 to CE-1, PE-2 acts as the SRv6 ingress PE node, while PE-1 acts as the SRv6 egress PE node. To explain the SRv6 encapsulation concept, the topology does not need an SRv6 transit router because SRv6 transit routers simply forward SRv6-encapsulated packets via IPv6 route table lookup without any other processing.

SRv6 and FPE are configured only on PE-1 and on PE-2. RR-3 acts as the BGP RR in the control plane and does not participate in the SRv6 data transport that only exists between PE-1 and PE-2.

The **ping** and **traceroute** commands between IPv4 and IPv6 system addresses of CE-1 and CE-2 simulate data transport.

The configuration for this example topology is completely symmetrical. All **configure** and **show** command outputs for PE-1 also apply for PE-2, and similar for CE-1 and CE-2.

The following sections describe the configuration steps needed to establish SRv6 Encapsulation in the base routing instance.

Configure the transport network:

This configuration includes:

- ports and IPv6-only interfaces on PE-1, PE-2, and RR-3
- port cross connect (PXC) and FPE on PE-1 and PE-2 (using internal loopbacks on an FP4 MAC chip), as described in the Segment Routing over IPv6 chapter .
- IS-IS on PE-1, PE-2, and RR-3 including:
  - level 2 capability with wide metrics (for the 128-bit identifiers)
  - native IPv6 routing
  - as best practice, include also: **traffic-engineering** and **traffic-engineering-options** on PE-1 and PE-2
  - advertise the router capability within the autonomous system (AS) (not for RR-3)
- BGP on PE-1, PE-2, and RR-3, with internal group “gr_v6_internal” that includes:
  - IPv4 and IPv6 address families
  - **extended-nh-encoding** for IPv4
  - **advertise-ipv6-next-hops** for IPv4
  - **next-hop-self** (not for RR-3)
  - BGP neighbor **system** IPv6 addresses

The core network topology uses IPv6 for BGP peering (with 16-byte next hop addresses), so to advertise and receive IPv4 routes (which have 4-byte next hop addresses) with IPv6 next hop addresses the commands **advertise-ipv6-next-hops** and **extended-nh-encoding** need to be configured at the BGP, group, or neighbor level. The **advertise-ipv6-next-hops** command instructs the system to advertise IPv4 routes with an IPv6 next hop address. The **extended-nh-encoding** command configures BGP to advertise the capability to receive IPv4 routes with an IPv6 next hop address.

The following example configuration applies for PE-1 and is similar for PE-2.

The following example configuration applies for RR-3:

```
[/]
A:admin@PE-1# configure {
    router "Base" {
        autonomous-system 64500
        interface "int-PE-1-PE-2" {
            description "interface between PE-1 and PE-2"
            port 1/1/c1/1:1000
            ipv6 {
                address 2001:db8::168:12:1 {
                    prefix-length 126
                }
            }
        }
        interface "int-PE-1-RR-3" {
            description "interface between PE-1 and RR-3"
            port 1/1/c2/1:1000
            ipv6 {
```

```
                address 2001:db8::168:13:1 {
                    prefix-length 126
                }
            }
        }
        interface "system" {
            description "system interface of PE-1"
            ipv6 {
                address 2001:db8::2:1 {
                    prefix-length 128
                }
            }
        }
        isis 0 {
            admin-state enable
            advertise-router-capability as
            ipv6-routing native
            level-capability 2    # required for SRv6
            router-id 1.1.1.1
            traffic-engineering true
            area-address [49.0001]
            traffic-engineering-options {
                ipv6 true
                application-link-attributes {
                }
            }
            interface "int-PE-1-PE-2" {
                interface-type point-to-point
            }
            interface "int-PE-1-RR-3" {
                interface-type point-to-point
            }
            interface "system" {
                passive true
            }
            level 2 {
                wide-metrics-only true    # required for SRv6
            }
        }
        bgp {
            min-route-advertisement 1
            router-id 2.2.2.1
            rapid-withdrawal true
            split-horizon true
            ebgp-default-reject-policy {
                import false    # do not refuse eBGP imported policies
                export false    # do not prevent eBGP exported policies
            }
            group "gr_v6_internal" {
                description "internal bgp group on PE-1"
                next-hop-self true
                type internal
                family {
                    ipv4 true
                    ipv6 true
                }
                extended-nh-encoding {
                    ipv4 true
                }
                advertise-ipv6-next-hops {
                    ipv4 true
                }
            }
            neighbor "2001:db8::2:3" {    # RR-3 system address
```

```
                group "gr_v6_internal"
            }
        }
    exit all
```

The following example configuration applies for RR-3:

```
[/]
A:admin@RR-3# configure {
    router "Base" {
        autonomous-system 64500
        interface "int-RR-3-PE-1" {
            description "interface between RR-3 and PE-1"
            port 1/1/c1/1:1000
            ipv6 {
                address 2001:db8::168:13:2 {
                    prefix-length 126
                }
            }
        }
        interface "int-RR-3-PE-2" {
            description "interface between RR-3 and PE-2"
            port 1/1/c2/1:1000
            ipv6 {
                address 2001:db8::168:23:2 {
                    prefix-length 126
                }
            }
        }
        interface "system" {
            description "system interface of RR-3"
            ipv6 {
                address 2001:db8::2:3 {
                    prefix-length 128
                }
            }
        }
        isis 0 {
            admin-state enable
            ipv6-routing native
            level-capability 2    # required for SRv6
            router-id 1.1.1.3
            area-address [49.0001]
            interface "int-RR-3-PE-1" {
                interface-type point-to-point
            }
            interface "int-RR-3-PE-2" {
                interface-type point-to-point
            }
            interface "system" {
                passive true
            }
            level 2 {
                wide-metrics-only true    # required for SRv6
            }
        }
        bgp {
            min-route-advertisement 1
            router-id 2.2.2.3
            rapid-withdrawal true
            split-horizon true
            group "gr_v6_internal" {
                description "internal bgp group on RR-3"
```

```
                type internal
                family {
                    ipv4 true
                    ipv6 true
                }
                cluster {
                    cluster-id 3.3.3.3
                }
                extended-nh-encoding {
                    ipv4 true
                }
                advertise-ipv6-next-hops {
                    ipv4 true
                }
            }
            neighbor "2001:db8::2:1" {    # PE-1 system address
                group "gr_v6_internal"
            }
            neighbor "2001:db8::2:2" {    # PE-2 system address
                group "gr_v6_internal"
            }
        }
    exit all
```

Configure CE-1 and CE-2 for native IPv4 and IPv6 data

This configuration includes:

- ports and IPv4 and IPv6 interfaces between CE-1 and PE-1 and between CE-2 and PE-2
- an IPv4 system address and an IPv6 system address for CE-1 and for CE-2
- BGP, with external group “gr_v6_external” that includes the following capabilities:
  - IPv4 and IPv6 address families
  - `extended-nh-encoding` for IPv4
  - `advertise-ipv6-next-hops` for IPv4
  - BGP neighbor **interface** IPv6 addresses, with BGP neighbors in a different external autonomous system

The following example configuration applies for PE-1 and is similar for PE-2. The **strip-srv6-tlvs** command (per address family) prevents PE-1 from advertising SRv6 TLVs to the BGP neighbor.

```
[/]
A:admin@PE-1# configure {
    router "Base" {
        interface "int-PE-1-CE-1" {
            description "interface between PE-1 and CE-1"
            port 1/1/c6/1:1000
            ipv4 {
                primary {
                    address 172.16.14.1
                    prefix-length 30
                }
            }
            ipv6 {
                address 2001:db8::168:14:1 {
                    prefix-length 126
                }
```

```
            }
        }
        bgp {
            group "gr_v6_external" {
                description "external bgp group on PE-1"
                family {
                    ipv4 true
                    ipv6 true
                }
                extended-nh-encoding {
                    ipv4 true
                }
                advertise-ipv6-next-hops {
                    ipv4 true
                }
            }
            neighbor "2001:db8::168:14:2" {
                group "gr_v6_external"
                type external
                peer-as 64504
                segment-routing-v6 {
                    route-advertisement {
                        family ipv4 {
                            strip-srv6-tlvs true
                        }
                        family ipv6 {
                            strip-srv6-tlvs true
                        }
                    }
                }
            }
        }
    exit all
```

The following example configuration applies for CE-2 and is similar for CE-1.

```
[/]
A:admin@CE-2# configure {
    router "Base" {
        autonomous-system 64505
        interface "int-CE-2-PE-2" {
            description "interface between CE-2 and PE-2"
            port 1/1/c1/1:1000
            ipv4 {
                primary {
                    address 172.16.25.2
                    prefix-length 30
                }
            }
            ipv6 {
                address 2001:db8::168:25:2 {
                    prefix-length 126
                }
            }
        }
        interface "system" {
            description "system interface of CE-2"
            ipv4 {
                primary {
                    address 192.0.2.5    # used for IPv4 ping
                    prefix-length 32
                }
            }
```

```
            ipv6 {
                address 2001:db8::2:5 {    # used for IPv6 ping
                    prefix-length 128
                }
            }
        }
        bgp {
            min-route-advertisement 1
            router-id 2.2.2.5
            rapid-withdrawal true
            split-horizon true
            ebgp-default-reject-policy {
                import false    # do not refuse eBGP imported policies
                export false    # do not prevent eBGP exported policies
            }
            group "gr_v6_external" {
                description "external bgp group on CE-2"
                family {
                    ipv4 true
                    ipv6 true
                }
                extended-nh-encoding {
                    ipv4 true
                }
                advertise-ipv6-next-hops {
                    ipv4 true
                }
            }
            neighbor "2001:db8::168:25:1" {
                group "gr_v6_external"
                type external
                peer-as 64500
            }
        }
    exit all
```

Ensure the export of the system addresses of CE-1 and CE-2

Configure a policy on CE-2 that imports the IPv4 and IPv6 prefixes into BGP. Configure a similar policy on CE-1.

```
[/]
A:admin@CE-2# configure {
    policy-options {
        prefix-list "CE-2_prefixes" {
            prefix 192.0.2.5/32 type exact {
            }
            prefix 2001:db8::2:5/128 type exact {
            }
        }
        policy-statement "policy-export-bgp" {
            entry 10 {
                from {
                    prefix-list ["CE-2_prefixes"]
                }
                action {
                    action-type accept
                }
            }
        }
```

```
    exit all
```

Apply this policy on CE-2 to the BGP neighbor PE-2. Perform a similar configuration on CE-1 to the BGP neighbor PE-1.

```
[ex:/configure router "Base"]
A:admin@CE-2#
        bgp {
            neighbor "2001:db8::168:25:1" {
                export {
                    policy ["policy-export-bgp"]
                }
            }
        exit all
```

Verify the IPv4 and IPv6 route tables. The corresponding FIBs can be verified with the **show router fib 1 ipv4** and **show router fib 1 ipv6** commands.

On CE-1:

192.0.2.4/32 is the IPv4 system address of CE-1. 192.0.2.5/32 is the IPv4 system address of CE-2, which is reached via PE-1.

```
[/]
A:admin@CE-1# show router route-table ipv4

===============================================================================
Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
172.16.14.0/30                                Local   Local     00h05m19s  0
       int-CE-1-PE-1                                                0
192.0.2.4/32                                  Local   Local     00h05m19s  0
       system                                                       0
192.0.2.5/32                                  Remote  BGP       00h00m26s  170
       2001:db8::168:14:1                                           0
-------------------------------------------------------------------------------
No. of Routes: 3
---snip---
===============================================================================
```

2001:db8::2:4/128 is the IPv6 system address of CE-1. 2001:db8::2:5/128 is the IPv6 system address of CE-2, which is reached via PE-1.

```
[/]
A:admin@CE-1# show router route-table ipv6

===============================================================================
IPv6 Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
2001:db8::2:4/128                             Local   Local     00h05m19s  0
       system                                                       0
2001:db8::2:5/128                             Remote  BGP       00h00m26s  170
       2001:db8::168:14:1                                           0
2001:db8::168:14:0/126                        Local   Local     00h05m18s  0
       int-CE-1-PE-1                                                0
```

```
-------------------------------------------------------------------------------
No. of Routes: 3
---snip---
===============================================================================
```

On PE-1:

```
[/]
A:admin@PE-1# show router route-table ipv4

===============================================================================
Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                              Type    Proto     Age        Pref
      Next Hop[Interface Name]                                      Metric
-------------------------------------------------------------------------------
172.16.14.0/30                                  Local   Local     00h06m21s  0
       int-PE-1-CE-1                                                  0
192.0.2.4/32                                    Remote  BGP       00h00m40s  170
       2001:db8::168:14:2                                             0
192.0.2.5/32                                    Remote  BGP       00h00m21s  170
       fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                           10
-------------------------------------------------------------------------------
No. of Routes: 3
---snip---
===============================================================================
```

```
[/]
A:admin@PE-1# show router route-table ipv6

===============================================================================
IPv6 Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                              Type    Proto     Age        Pref
      Next Hop[Interface Name]                                      Metric
-------------------------------------------------------------------------------
2001:db8::2:1/128                               Local   Local     00h12m43s  0
       system                                                         0
2001:db8::2:2/128                               Remote  ISIS      00h11m03s  18
       fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                           10
2001:db8::2:3/128                               Remote  ISIS      00h10m52s  18
       fe80::612:1ff:fe01:1-"int-PE-1-RR-3"                           10
2001:db8::2:4/128                               Remote  BGP       00h00m40s  170
       2001:db8::168:14:2                                             0
2001:db8::2:5/128                               Remote  BGP       00h00m21s  170
       fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                           10
2001:db8::168:12:0/126                          Local   Local     00h12m43s  0
       int-PE-1-PE-2                                                  0
2001:db8::168:13:0/126                          Local   Local     00h12m42s  0
       int-PE-1-RR-3                                                  0
2001:db8::168:14:0/126                          Local   Local     00h06m20s  0
       int-PE-1-CE-1                                                  0
2001:db8::168:23:0/126                          Remote  ISIS      00h11m03s  18
       fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                           20
-------------------------------------------------------------------------------
No. of Routes: 9
---snip---
===============================================================================
```

On PE-2:

```
[/]
```

```
A:admin@PE-2# show router route-table ipv4

===============================================================================
Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                              Type    Proto     Age        Pref
      Next Hop[Interface Name]                                      Metric
-------------------------------------------------------------------------------
172.16.25.0/30                                  Local   Local     00h06m04s  0
       int-PE-2-CE-2                                                0
192.0.2.4/32                                    Remote  BGP       00h00m42s  170
       fe80::60a:1ff:fe01:1-"int-PE-2-PE-1"                         10
192.0.2.5/32                                    Remote  BGP       00h00m24s  170
       2001:db8::168:25:2                                           0
-------------------------------------------------------------------------------
No. of Routes: 3
---snip---
===============================================================================
```

```
[/]
A:admin@PE-2# show router route-table ipv6

===============================================================================
IPv6 Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                              Type    Proto     Age        Pref
      Next Hop[Interface Name]                                      Metric
-------------------------------------------------------------------------------
2001:db8::2:1/128                               Remote  ISIS      00h11m00s  18
       fe80::60a:1ff:fe01:1-"int-PE-2-PE-1"                         10
2001:db8::2:2/128                               Local   Local     00h12m27s  0
       system                                                       0
2001:db8::2:3/128                               Remote  ISIS      00h10m54s  18
       fe80::612:1ff:fe01:b-"int-PE-2-RR-3"                         10
2001:db8::2:4/128                               Remote  BGP       00h00m42s  170
       fe80::60a:1ff:fe01:1-"int-PE-2-PE-1"                         10
2001:db8::2:5/128                               Remote  BGP       00h00m24s  170
       2001:db8::168:25:2                                           0
2001:db8::168:12:0/126                          Local   Local     00h12m26s  0
       int-PE-2-PE-1                                                0
2001:db8::168:13:0/126                          Remote  ISIS      00h11m00s  18
       fe80::60a:1ff:fe01:1-"int-PE-2-PE-1"                         20
2001:db8::168:23:0/126                          Local   Local     00h12m26s  0
       int-PE-2-RR-3                                                0
2001:db8::168:25:0/126                          Local   Local     00h06m03s  0
       int-PE-2-CE-2                                                0
-------------------------------------------------------------------------------
No. of Routes: 9
---snip---
===============================================================================
```

IPv4 data transport is not possible between CE-1 and CE-2. Verify this with a **ping** from CE-1 to the IPv4 system address that CE-2 advertises.

```
[/]
A:admin@CE-1# ping 192.0.2.5
PING 192.0.2.5 56 data bytes
...   ...   ...   ...   . Request timed out. icmp_seq=1.
Request timed out. icmp_seq=2.
---snip---
---- 192.0.2.5 PING Statistics ----
5 packets transmitted, 0 packets received, 100% packet loss
```

IPv6 data transport is possible between CE-1 and CE-2, although not by using SRv6 between PE-1 and PE-2 but by using native IPv6. Verify this with a **ping** and a **traceroute** from CE-1 to the IPv6 system address that CE-2 advertises. Native IPv6 data flows over an IPv6 interface from CE-1 to PE-1, from there over an IPv6 interface to PE-2, and from there over an IPv6 interface to CE-2. The same is true for data transport between CE-2 and CE-1.

```
[/]
A:admin@CE-1# ping 2001:db8::2:5
PING 2001:db8::2:5 56 data bytes
64 bytes from 2001:db8::2:5 icmp_seq=1 hlim=62 time=2.18ms.
---snip---
---- 2001:db8::2:5 PING Statistics ----
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min = 1.85ms, avg = 2.07ms, max = 2.18ms, stddev = 0.122ms
```

```
[/]
A:admin@CE-1# traceroute 2001:db8::2:5
traceroute to 2001:db8::2:5, 30 hops max, 60 byte packets
  1  2001:db8::168:14:1 (2001:db8::168:14:1)    1.02 ms  0.790 ms  0.837 ms
  2  2001:db8::168:12:2 (2001:db8::168:12:2)    1.28 ms  1.26 ms  1.41 ms
  3  2001:db8::2:5 (2001:db8::2:5)    1.80 ms  1.99 ms  1.82 ms
```

Configure SRv6 in the router Base context on PE-1 and PE-2

Configure the locator in the **router Base segment-routing segment-routing-v6** context on PE-2. Perform a similar configuration on PE-1, with **ip-prefix** 2001:db8:aaaa:101::/64 for locator “PE-1_loc”.

```
[ex:/configure router "Base" segment-routing]
A:admin@PE-2#
            segment-routing-v6 {
                locator "PE-2_loc" {
                    admin-state enable
                    block-length 48
                    prefix {
                        ip-prefix 2001:db8:aaaa:102::/64
                    }
                }
            exit all
```

Configure the FPEs on PE-1 and PE-2.

```
[ex:/configure]
A:admin@PE-2#
    fwd-path-ext {
        fpe 1 {
            path {
                pxc 1
            }
            application {
                srv6 {
                    type origination
                }
            }
        }
        fpe 2 {
            path {
                pxc 2
```

```
            }
            application {
                srv6 {
                    type termination
                }
            }
        }
    exit all
```

Use FPE 1 as the SRv6 origination FPE in the **router Base segment-routing segment-routing-v6** context and FPE 2 as the SRv6 termination FPE in the **router Base segment-routing segment-routing-v6 locator** context on PE-2. Perform a similar configuration on PE-1 for locator “PE-1_loc”.

```
[ex:/configure router "Base" segment-routing]
A:admin@PE-2#
            segment-routing-v6 {
                origination-fpe [1]
                locator "PE-2_loc" {
                    admin-state enable
                    termination-fpe [2]
                }
            exit all
```

Configure the SRv6 End function (equivalent to an IPv4 node SID) and SRv6 End.X functions (equivalent to IPv4 Adjacency SIDs) in the **router Base segment-routing segment-routing-v6 base-routing-instance locator** context on PE-2. Perform a similar configuration on PE-1 for locator “PE-1_loc”.

```
[ex:/configure router "Base" segment-routing]
A:admin@PE-2#
            segment-routing-v6 {
                base-routing-instance {
                    locator "PE-2_loc" {
                        function {
                            end 1 {
                                srh-mode usp
                            }
                            end-x-auto-allocate psp protection unprotected { }
                        }
                    }
                }
            exit all
```

Advertise the locator in IS-IS while ensuring level 2 capability on PE-2. Perform a similar configuration on PE-1 for locator “PE-1_loc”.

```
[ex:/configure router "Base" isis 0]
A:admin@PE-2#
            segment-routing-v6 {
                admin-state enable
                locator "PE-2_loc" {
                    level-capability 2
                }
            exit all
```

A summary on locator and origination FPE configuration can be verified with the **show router segment-routing-v6 summary** command.

Verify the SRv6 local SIDs on PE-2 and similar on PE-1. Three SRv6 local SIDs are created: one for the statically configured SRv6 End function (configured in the base context) and two for the auto-allocated

SRv6 End.X functions (one facing PE-1 and one facing RR-3). All three SRv6 local SIDs are concatenated with the locator. The statically configured SRv6 End function appears first with function number 1. The auto-allocated SRv6 End.X functions get subsequent function numbers, 2 and 4 respectively. RR-3 has no SRv6 configuration and does not have these SRv6 local SIDs and SRv6 functions.

```
[/]
A:admin@PE-2# show router segment-routing-v6 local-sid

===============================================================================
Segment Routing v6 Local SIDs
===============================================================================
SID                                                 Type           Function
  Locator
  Context
-------------------------------------------------------------------------------
2001:db8:aaaa:102:0:1000::                          End            1
  PE-2_loc
  Base
2001:db8:aaaa:102:0:2000::                          End.X          2
  PE-2_loc
    None
2001:db8:aaaa:102:0:4000::                          End.X          4
  PE-2_loc
    None
-------------------------------------------------------------------------------
SIDs : 3
-------------------------------------------------------------------------------
===============================================================================
```

Verify the SRv6 base routing instance details on PE-2 and similar on PE-1. The SRv6 End function is statically configured. There is an auto-allocated SRv6 End.X function for each IS-IS neighbor.

```
[/]
A:admin@PE-2# show router segment-routing-v6 base-routing-instance

===============================================================================
Segment Routing v6 Base Routing Instance
===============================================================================
Locator
  Type       Function    SID                                      Status/InstId
    SRH-mode Protection  Interface
-------------------------------------------------------------------------------
PE-2_loc
  End                 1 2001:db8:aaaa:102:0:1000::              ok
    USP
-------------------------------------------------------------------------------
  Auto-allocated End.X: PSP Unprotected,
-------------------------------------------------------------------------------
  End.X              *2 2001:db8:aaaa:102:0:2000::              0
    PSP      Unprotected int-PE-2-PE-1
    ISIS Level: L2 Mac Address: 04:0a:01:01:00:01 Nbr Sys Id: 0010.0100.1001
  End.X              *4 2001:db8:aaaa:102:0:4000::              0
    PSP      Unprotected int-PE-2-RR-3
    ISIS Level: L2 Mac Address: 04:12:01:01:00:0b Nbr Sys Id: 0010.0100.1003
-------------------------------------------------------------------------------
===============================================================================
Legend: * - System allocated
```

Verify the IPv6 route table on PE-1. The IPv6 route table has also routes to the local and the learned remote locators and to the local SRv6 function SIDs. The remotely configured locator prefix of PE-2 is

reached via an SRv6 tunnel. The routes with protocol “SRv6” correspond with the locally configured locator prefix of PE-1, or the locally configured SRv6 End function.

```
[/]
A:admin@PE-1# show router route-table ipv6

===============================================================================
IPv6 Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
---snip---
2001:db8:aaaa:101::/64                        Local   SRV6      00h02m44s  3
       fe80::201-"_tmnx_fpe_2.a"                                      0
2001:db8:aaaa:101:0:1000::/128                Local   SRV6      00h01m33s  3
       Black Hole                                                     0
2001:db8:aaaa:101:0:2000::/128                Local   ISIS      00h00m45s  18
       fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                           10
2001:db8:aaaa:101:0:4000::/128                Local   ISIS      00h00m45s  18
       fe80::612:1ff:fe01:1-"int-PE-1-RR-3"                           10
2001:db8:aaaa:102::/64                        Remote  ISIS      00h00m31s  18
       2001:db8:aaaa:102::/64 (tunneled:SRV6-ISIS)                    10
-------------------------------------------------------------------------------
No. of Routes: 14
---snip---
===============================================================================
```

Verify the IS-IS routes on PE-1 and similar on PE-2. This corresponds with the information in the route table (and FIB). IS-IS is not configured on CE-1 and CE-2, so CE-1 and CE-2 have no IS-IS routes.

```
[/]
A:admin@PE-1# show router isis routes

===============================================================================
Rtr Base ISIS Instance 0 Route Table
===============================================================================
Prefix[Flags]                     Metric     Lvl/Typ     Ver.  SysID/Hostname
  NextHop                                                MT     AdminTag/SID[F]
-------------------------------------------------------------------------------
2001:db8::2:1/128                 0          2/Int.      2     PE-1
   ::                                                      0       0
2001:db8::2:2/128                 10         2/Int.      9     PE-2
   fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                    0       0
2001:db8::2:3/128                 10         2/Int.      9     RR-3
   fe80::612:1ff:fe01:1-"int-PE-1-RR-3"                    0       0
2001:db8::168:12:0/126            10         2/Int.      4     PE-1
   ::                                                      0       0
2001:db8::168:13:0/126            10         2/Int.      4     PE-1
   ::                                                      0       0
2001:db8::168:23:0/126            20         2/Int.      9     PE-2
   fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                    0       0
2001:db8:aaaa:101::/64            0          2/Int.      11    PE-1
   ::                                                      0       0
2001:db8:aaaa:102::/64            10         2/Int.      10    PE-2
   fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                    0       0
-------------------------------------------------------------------------------
No. of Routes: 8 (8 paths)
---snip---
===============================================================================
```

The locator prefixes and who advertises them can be verified with the **show router isis segment-routing-v6 locator** command. The SRv6 End SIDs and who advertises them can be verified with the **show router isis segment-routing-v6 end-sid** command.

The IS-IS data base can be verified with the **show router isis database detail** command.

The output of this command provides information on each IS-IS-enabled router. Per uniquely identified IS-IS-enabled router, the SRv6 information indicates:

- the IS-IS-advertised router capabilities
- the advertised SRv6 locator TLV
- the advertised configured SRv6 End SID and auto-allocated SRv6 End.X SIDs

The BGP groups can be verified with the **show router bgp group** command. PE-1 and PE-2 know the iBGP and eBGP peers. RR-3 only knows the iBGP peers. CE-1 and CE-2 only know the eBGP peers.

The BGP next hops can be verified with the **show router bgp next-hop ipv4** and **show router bgp next-hop ipv6** commands.

Configure SRv6 End.DT4 and SRv6 End.DT6 functions on PE-1 and PE-2

Configure SRv6 End.DT4 and SRv6 End.DT6 functions in the **router Base segment-routing segment-routing-v6 base-routing-instance locator function** context on PE-1. They can have statically or automatically allocated values. For statically allocated values, an SRv6 reserved label block must be configured. Perform an identical configuration on PE-2.

```
[ex:/configure router "Base"]
A:admin@PE-1#
        mpls-labels {
            sr-labels {
                start 20000
                end 20999
            }
            reserved-label-block "SRv6" {
                start-label 30100
                end-label 30199
            }
        exit all
```

This SRv6 reserved label block must be referenced in the **router Base segment-routing segment-routing-v6 locator static-function** context on PE-2, where also the total number of static functions, including the already existing SRv6 End function (with value 1), must be set. Perform a similar configuration on PE-1 for locator “PE-1_loc”.

```
[ex:/configure router "Base" segment-routing]
A:admin@PE-2#
            segment-routing-v6 {
                source-address 2001:db8::2:2
                locator "PE-2_loc" {
                    admin-state enable
                    static-function {
                        max-entries 3
                        label-block "SRv6"
                    }
                }
                base-routing-instance {
                    locator "PE-2_loc" {
```

```
                        function {
                            end-dt4 {
                                value 2
                            }
                            end-dt6 {
                                value 3
                            }
                        }
                    }
                }
            exit all
```

The SRv6 End.DT4 and SRv6 End.DT6 functions are allocated the unique static values of 2 and 3 respectively. The values allocated must not exceed the **max-entries** value.

Each PE must resolve the BGP next hop to an SRv6 End.DT4 or End.DT6 Segment ID. Therefore, each PE must advertise route prefixes within a BGP update message that includes an SRv6 Services TLV. This is achieved by configuring the **add-srv6-tlvs** command along with the locator value for each address family, IPv4 and IPv6.

When a PE receives a BGP update that includes the SRv6 Services TLV, the default behavior is to ignore this TLV, and resolve the next hop to the tunnel type configured in an **auto-bind-tunnel** statement. To override this behavior, **ignore-received-srv6-tlvs** must be set to false for IPv4 and IPv6 address families on PE-1. Perform a similar configuration on PE-2 for locator “PE-2_loc”.

```
[ex:/configure router "Base" bgp]
A:admin@PE-1#
            segment-routing-v6 {
                family ipv4 {
                    ignore-received-srv6-tlvs false
                    add-srv6-tlvs {
                        locator-name "PE-1_loc"
                    }
                }
                family ipv6 {
                    ignore-received-srv6-tlvs false
                    add-srv6-tlvs {
                        locator-name "PE-1_loc"
                    }
                }
            exit all
```

CE-2 sends BGP updates to PE-2 for the IPv4 and the IPv6 address families respectively. Each BGP update advertises the IPv4 or IPv6 address family, the reachable network prefixes, and the autonomous system to which they belong. PE-2 adds an SRv6 Services TLV, indicating that resolution to an SRv6 SID is available, making use of the endpoint behavior that is configured for the IPv4 or IPv6 address family on the locator. PE-2 advertises the BGP updates to PE-1 via the RR. PE-1 programs the route prefixes in its route table and FIB with an SRv6 tunnel next hop, and forwards the BGP updates to CE-1. CE-1 programs the learned route prefixes in its route table and FIB.

Similar BGP updates flow from CE-1 to CE-2, via PE-1, RR-3, and PE-2. PE-1 and PE-2 advertise only the SRv6 SIDs for the SRv6 End.DT4 and SRv6 End.DT6 functions.

After the BGP updates, the IS-IS data base remains the same, except for the renumbering of the SRv6 End.X functions. This can be verified with the **show router isis database detail** command.

The BGP next hops remain the same, except for the next hops to the system addresses of PE-1 and PE-2 that switch owner from “ISIS” to “N/A”. This can be verified with the **show router bgp next-hop ipv4** and **show router bgp next-hop ipv6** commands.

When debug logging for BGP updates is configured, the configuration results in the following BGP update logs for the IPv4 address family.

Focus as an example on prefix 192.0.2.5/32 and on prefix 192.0.2.4.1/32, but in the other direction.

Verify the IPv4 BGP routes.

CE-2 advertises route prefix 192.0.2.5/32 to PE-2 (in RIB Out Entries).

```
[/]
A:admin@CE-2# show router bgp routes 192.0.2.5 hunt
===============================================================================
 BGP Router ID:2.2.2.5          AS:64505       Local AS:64505
===============================================================================
---snip---
===============================================================================
BGP IPv4 Routes
===============================================================================
-------------------------------------------------------------------------------
RIB In Entries
-------------------------------------------------------------------------------

-------------------------------------------------------------------------------
RIB Out Entries
-------------------------------------------------------------------------------
Network        : 192.0.2.5/32
Nexthop        : 2001:db8::168:25:2
Path Id        : None
To             : 2001:db8::168:25:1
Res. Protocol  : INVALID               Res. Metric    : 0
Res. Nexthop   : n/a
Local Pref.    : n/a                   Interface Name : NotAvailable
Aggregator AS  : None                  Aggregator     : None
Atomic Aggr.   : Not Atomic            MED            : None
AIGP Metric    : None                  IGP Cost       : n/a
Connector      : None
Community      : No Community Members
Cluster        : No Cluster Members
Originator Id  : None                  Peer Router Id : 2.2.2.2
Origin         : IGP
AS-Path        : 64505
Route Tag      : 0
Neighbor-AS    : 64505
Orig Validation: NotFound
Source Class   : 0                     Dest Class     : 0

-------------------------------------------------------------------------------
Routes : 1
===============================================================================
```

PE-2 receives the BGP update which CE-2 sends for the IPv4 address family:

```
[/]
A:admin@PE-2# show log log-id "log_2"

---snip---
5 2022/07/19 11:29:41.236 CEST MINOR: DEBUG #2001 Base Peer 1: 2001:db8::168:25:2
"Peer 1: 2001:db8::168:25:2: UPDATE
Peer 1: 2001:db8::168:25:2 - Received BGP UPDATE:
    Withdrawn Length = 0
    Total Path Attr Length = 43
    Flag: 0x90 Type: 14 Len: 26 Multiprotocol Reachable NLRI:
        Address Family IPV4
```

```
        NextHop len 16 Global NextHop 2001:db8::168:25:2
        192.0.2.5/32
    Flag: 0x40 Type: 1 Len: 1 Origin: 0
    Flag: 0x40 Type: 2 Len: 6 AS Path:
        Type: 2 Len: 1 < 64505 >
"
---snip---
```

Upon receipt of the BGP update from CE-2, PE-2 programs route prefix 192.0.2.5/32 in its route table and FIB, with the interface towards CE-2 as next hop.

Verify the resulting IPv4 route table on PE-2:

```
[/]
A:admin@PE-2# show router route-table ipv4

===============================================================================
Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                              Type    Proto     Age        Pref
      Next Hop[Interface Name]                                      Metric
-------------------------------------------------------------------------------
172.16.25.0/30                                  Local   Local     00h16m25s  0
       int-PE-2-CE-2                                                 0
192.0.2.4/32                                    Remote  BGP       00h00m49s  170
       2001:db8:aaaa:101:0:2000:: (tunneled:SRV6)                    10
192.0.2.5/32                                    Remote  BGP       00h10m45s  170
       2001:db8::168:25:2                                            0
-------------------------------------------------------------------------------
No. of Routes: 3
---snip---
===============================================================================
```

Verify the corresponding IPv4 BGP routes on PE-2:

```
[/]
A:admin@PE-2# show router bgp routes 192.0.2.5 hunt
===============================================================================
 BGP Router ID:2.2.2.2          AS:64500       Local AS:64500
===============================================================================
---snip---
===============================================================================
BGP IPv4 Routes
===============================================================================
-------------------------------------------------------------------------------
RIB In Entries
-------------------------------------------------------------------------------
Network        : 192.0.2.5/32
Nexthop        : 2001:db8::168:25:2
Path Id        : None
From           : 2001:db8::168:25:2
Res. Protocol  : LOCAL                  Res. Metric    : 0
Res. Nexthop   : 2001:db8::168:25:2
Local Pref.    : None                   Interface Name : int-PE-2-CE-2
Aggregator AS  : None                   Aggregator     : None
Atomic Aggr.   : Not Atomic             MED            : None
AIGP Metric    : None                   IGP Cost       : 0
Connector      : None
Community      : No Community Members
Cluster        : No Cluster Members
Originator Id  : None                   Peer Router Id : 2.2.2.5
Fwd Class      : None                   Priority       : None
Flags          : Used Valid Best IGP In-RTM
```

```
Route Source   : External
AS-Path        : 64505
Route Tag      : 0
Neighbor-AS    : 64505
Orig Validation: NotFound
Source Class   : 0                     Dest Class     : 0
Add Paths Send : Default
RIB Priority   : Normal
Last Modified  : 00h10m45s

-------------------------------------------------------------------------------
RIB Out Entries
-------------------------------------------------------------------------------
Network        : 192.0.2.5/32
Nexthop        : 2001:db8::2:2
Path Id        : None
To             : 2001:db8::2:3
Res. Protocol  : INVALID               Res. Metric    : 0
Res. Nexthop   : n/a
Local Pref.    : 100                   Interface Name : NotAvailable
Aggregator AS  : None                  Aggregator     : None
Atomic Aggr.   : Not Atomic            MED            : None
AIGP Metric    : None                  IGP Cost       : 0
Connector      : None
Community      : No Community Members
Cluster        : No Cluster Members
Originator Id  : None                  Peer Router Id : 2.2.2.3
Origin         : IGP
AS-Path        : 64505
Route Tag      : 0
Neighbor-AS    : 64505
Orig Validation: NotFound
Source Class   : 0                     Dest Class     : 0
SRv6 TLV Type  : SRv6 L3 Service TLV (5)
SRv6 SubTLV    : SRv6 SID Information (1)
Sid            : 2001:db8:aaaa:102:0:2000::
Behavior       : End.DT4 (19)
SRv6 SubSubTLV : SRv6 SID Structure (1)
Loc-Block-Len  : 48                    Loc-Node-Len   : 16
Func-Len       : 20                    Arg-Len        : 0
Tpose-Len      : 0                     Tpose-offset   : 0

-------------------------------------------------------------------------------
Routes : 2
===============================================================================
```

PE-2 then advertises route prefix 192.0.2.5/32 (in RIB Out Entries), via RR-3, and inserts the SRv6 Services TLV. This TLV carries an SRv6 Service Information sub-TLV that contains the End.DT4 SID.

PE-1 receives (via RR-3) the BGP update which PE-2 sends for the IPv4 address family:

```
[/]
A:admin@PE-1# show log log-id "log_2"

---snip---
3 2022/07/19 11:39:38.404 CEST MINOR: DEBUG #2001 Base Peer 1: 2001:db8::2:3
"Peer 1: 2001:db8::2:3: UPDATE
Peer 1: 2001:db8::2:3 - Received BGP UPDATE:
    Withdrawn Length = 0
    Total Path Attr Length = 104
    Flag: 0x90 Type: 14 Len: 26 Multiprotocol Reachable NLRI:
        Address Family IPV4
        NextHop len 16 Global NextHop 2001:db8::2:2
```

```
        192.0.2.5/32
    Flag: 0x40 Type: 1 Len: 1 Origin: 0
    Flag: 0x40 Type: 2 Len: 6 AS Path:
        Type: 2 Len: 1 < 64505 >
    Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
    Flag: 0x80 Type: 9 Len: 4 Originator ID: 2.2.2.2
    Flag: 0x80 Type: 10 Len: 4 Cluster ID:
        3.3.3.3
    Flag: 0xc0 Type: 40 Len: 37 Prefix-SID-attr:
       SRv6 Services TLV (37 bytes):-
           Type: SRV6 L3 Service TLV (5)
           Length: 34 bytes, Reserved: 0x0
         SRv6 Service Information Sub-TLV (33 bytes)
            Type: 1 Len: 30 Rsvd1: 0x0
            SRv6 SID: 2001:db8:aaaa:102:0:2000::
            SID Flags: 0x0 Endpoint Behavior: 0x13 Rsvd2: 0x0
            SRv6 SID Sub-Sub-TLV
               Type: 1 Len: 6
               BL:48 NL:16 FL:20 AL:0 TL:0 TO:0
"
---snip---
```

Upon receipt of the BGP update from RR-3 on behalf of PE-2, PE-1 programs route prefix 192.0.2.5/32 in its route table and FIB. The presence of the SRv6 Services TLV indicates that the next hop is the SRv6 End.DT4 SID which, in turn, is resolved to the remote locator for PE-2.

PE-1 then advertises route prefix 192.0.2.5/32 to CE-1 (in RIB Out Entries).

Verify the resulting IPv4 route table on PE-1. The IPv4 route table has a route to the remote IPv4 system address of CE-2, via the End.DT4 SID of the remotely configured locator prefix of PE-2.

```
[/]
A:admin@PE-1# show router route-table ipv4

===============================================================================
Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
172.16.14.0/30                                Local   Local     00h16m41s  0
       int-PE-1-CE-1                                              0
192.0.2.4/32                                  Remote  BGP       00h11m01s  170
       2001:db8::168:14:2                                         0
192.0.2.5/32                                  Remote  BGP       00h00m46s  170
       2001:db8:aaaa:102:0:2000:: (tunneled:SRV6)                 10
-------------------------------------------------------------------------------
No. of Routes: 3
---snip---
===============================================================================
```

Verify the corresponding IPv4 BGP routes on PE-1:

```
[/]
A:admin@PE-1# show router bgp routes 192.0.2.5 hunt
===============================================================================
 BGP Router ID:2.2.2.1          AS:64500       Local AS:64500
===============================================================================
---snip---
===============================================================================
BGP IPv4 Routes
===============================================================================
```

```
-------------------------------------------------------------------------------
RIB In Entries
-------------------------------------------------------------------------------
Network        : 192.0.2.5/32
Nexthop        : 2001:db8::2:2
Path Id        : None
From           : 2001:db8::2:3
Res. Protocol  : ISIS                    Res. Metric    : 10
Res. Nexthop   : fe80::60e:1ff:fe01:1
Local Pref.    : 100                     Interface Name : int-PE-1-PE-2
Aggregator AS  : None                    Aggregator     : None
Atomic Aggr.   : Not Atomic              MED            : None
AIGP Metric    : None                    IGP Cost       : 10
Connector      : None
Community      : No Community Members
Cluster        : 3.3.3.3
Originator Id  : 2.2.2.2                 Peer Router Id : 2.2.2.3
Fwd Class      : None                    Priority       : None
Flags          : Used Valid Best IGP In-RTM
Route Source   : Internal
AS-Path        : 64505
Route Tag      : 0
Neighbor-AS    : 64505
Orig Validation: NotFound
Source Class   : 0                       Dest Class     : 0
Add Paths Send : Default
RIB Priority   : Normal
Last Modified  : 00h00m46s
SRv6 TLV Type  : SRv6 L3 Service TLV (5)
SRv6 SubTLV    : SRv6 SID Information (1)
Sid            : 2001:db8:aaaa:102:0:2000::
Behavior       : End.DT4 (19)
SRv6 SubSubTLV : SRv6 SID Structure (1)
Loc-Block-Len  : 48                      Loc-Node-Len   : 16
Func-Len       : 20                      Arg-Len        : 0
Tpose-Len      : 0                       Tpose-offset   : 0

-------------------------------------------------------------------------------
RIB Out Entries
-------------------------------------------------------------------------------
Network        : 192.0.2.5/32
Nexthop        : 2001:db8::168:14:1
Path Id        : None
To             : 2001:db8::168:14:2
Res. Protocol  : INVALID                 Res. Metric    : 0
Res. Nexthop   : n/a
Local Pref.    : n/a                     Interface Name : NotAvailable
Aggregator AS  : None                    Aggregator     : None
Atomic Aggr.   : Not Atomic              MED            : None
AIGP Metric    : None                    IGP Cost       : 10
Connector      : None
Community      : No Community Members
Cluster        : No Cluster Members
Originator Id  : None                    Peer Router Id : 2.2.2.4
Origin         : IGP
AS-Path        : 64500 64505
Route Tag      : 0
Neighbor-AS    : 64500
Orig Validation: NotFound
Source Class   : 0                       Dest Class     : 0

-------------------------------------------------------------------------------
Routes : 2
===============================================================================
```

CE-1 receives the BGP update which PE-1 sends for the IPv4 address family:

```
[/]
A:admin@CE-1# show log log-id "log_2"

---snip---
1 2022/07/19 11:39:39.220 CEST MINOR: DEBUG #2001 Base Peer 1: 2001:db8::168:14:1
"Peer 1: 2001:db8::168:14:1: UPDATE
Peer 1: 2001:db8::168:14:1 - Received BGP UPDATE:
    Withdrawn Length = 0
    Total Path Attr Length = 47
    Flag: 0x90 Type: 14 Len: 26 Multiprotocol Reachable NLRI:
        Address Family IPV4
        NextHop len 16 Global NextHop 2001:db8::168:14:1
        192.0.2.5/32
    Flag: 0x40 Type: 1 Len: 1 Origin: 0
    Flag: 0x40 Type: 2 Len: 10 AS Path:
        Type: 2 Len: 2 < 64500 64505 >
"
---snip---
```

Upon receipt of the BGP update from PE-1, CE-1 programs route prefix 192.0.2.5/32 in its route table and FIB with the interface towards PE-1 as the next hop.

Verify the resulting IPv4 route table on CE-1:

```
[/]
A:admin@CE-1# show router route-table ipv4

===============================================================================
Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
172.16.14.0/30                                Local   Local     00h15m40s  0
       int-CE-1-PE-1                                                0
192.0.2.4/32                                  Local   Local     00h15m40s  0
       system                                                       0
192.0.2.5/32                                  Remote  BGP       00h00m51s  170
       2001:db8::168:14:1                                           0
-------------------------------------------------------------------------------
No. of Routes: 3
---snip---
===============================================================================
```

Verify the corresponding IPv4 BGP routes on CE-1:

```
[/]
A:admin@CE-1# show router bgp routes 192.0.2.5 hunt
===============================================================================
 BGP Router ID:2.2.2.4          AS:64504       Local AS:64504
===============================================================================
---snip---
===============================================================================
BGP IPv4 Routes
===============================================================================
-------------------------------------------------------------------------------
RIB In Entries
-------------------------------------------------------------------------------
Network        : 192.0.2.5/32
Nexthop        : 2001:db8::168:14:1
```

```
Path Id         : None
From            : 2001:db8::168:14:1
Res. Protocol   : LOCAL                   Res. Metric    : 0
Res. Nexthop    : 2001:db8::168:14:1
Local Pref.     : None                    Interface Name : int-CE-1-PE-1
Aggregator AS   : None                    Aggregator     : None
Atomic Aggr.    : Not Atomic              MED            : None
AIGP Metric     : None                    IGP Cost       : 0
Connector       : None
Community       : No Community Members
Cluster         : No Cluster Members
Originator Id   : None                    Peer Router Id : 2.2.2.1
Fwd Class       : None                    Priority       : None
Flags           : Used Valid Best IGP In-RTM
Route Source    : External
AS-Path         : 64500 64505
Route Tag       : 0
Neighbor-AS     : 64500
Orig Validation: NotFound
Source Class    : 0                       Dest Class     : 0
Add Paths Send : Default
RIB Priority    : Normal
Last Modified   : 00h00m51s

-------------------------------------------------------------------------------
RIB Out Entries
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
Routes : 1
===============================================================================
```

Similar BGP update logs are generated for the IPv6 address family.

Focus as an example on prefix 2001:db8::2:5/128 and on prefix 2001:db8::2:4/128, but in the other direction.

Verify the IPv6 BGP routes.

CE-2 advertises route prefix 2001:db8::2:5/128 to PE-2 (in RIB Out Entries).

```
[/]
A:admin@CE-2# show router bgp routes 2001:db8::2:5/128 hunt
===============================================================================
 BGP Router ID:2.2.2.5          AS:64505       Local AS:64505
===============================================================================
---snip---
===============================================================================
BGP IPv6 Routes
===============================================================================
-------------------------------------------------------------------------------
RIB In Entries
-------------------------------------------------------------------------------

-------------------------------------------------------------------------------
RIB Out Entries
-------------------------------------------------------------------------------
Network         : 2001:db8::2:5/128
Nexthop         : 2001:db8::168:25:2
Path Id         : None
To              : 2001:db8::168:25:1
Res. Protocol   : INVALID                 Res. Metric    : 0
Res. Nexthop    : n/a
Local Pref.     : n/a                     Interface Name : NotAvailable
Aggregator AS   : None                    Aggregator     : None
```

```
Atomic Aggr.   : Not Atomic             MED            : None
AIGP Metric    : None                   IGP Cost       : n/a
Connector      : None
Community      : No Community Members
Cluster        : No Cluster Members
Originator Id  : None                   Peer Router Id : 2.2.2.2
Origin         : IGP
AS-Path        : 64505
Route Tag      : 0
Neighbor-AS    : 64505
Orig Validation: NotFound
Source Class   : 0                      Dest Class     : 0

-------------------------------------------------------------------------------
Routes : 1
===============================================================================
```

PE-2 receives the BGP update which CE-2 sends for the IPv6 address family:

```
[/]
A:admin@PE-2# show log log-id "log_2"

---snip---
6 2022/07/19 11:29:41.237 CEST MINOR: DEBUG #2001 Base Peer 1: 2001:db8::168:25:2
"Peer 1: 2001:db8::168:25:2: UPDATE
Peer 1: 2001:db8::168:25:2 - Received BGP UPDATE:
    Withdrawn Length = 0
    Total Path Attr Length = 55
    Flag: 0x90 Type: 14 Len: 38 Multiprotocol Reachable NLRI:
        Address Family IPV6
        NextHop len 16 Global NextHop 2001:db8::168:25:2
        2001:db8::2:5/128
    Flag: 0x40 Type: 1 Len: 1 Origin: 0
    Flag: 0x40 Type: 2 Len: 6 AS Path:
        Type: 2 Len: 1 < 64505 >
"
---snip---
```

Upon receipt of the BGP update from CE-2, PE-2 programs route prefix 2001:db8::2:5/128 in its route table and FIB, with the interface towards CE-2 as next hop.

Verify the resulting IPv6 route table on PE-2:

```
[/]
A:admin@PE-2# show router route-table ipv6

===============================================================================
IPv6 Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
---snip---
2001:db8::2:4/128                             Remote  BGP       00h00m49s  170
       2001:db8:aaaa:101:0:3000:: (tunneled:SRV6)                    10
2001:db8::2:5/128                             Remote  BGP       00h10m45s  170
       2001:db8::168:25:2                                            0
---snip---
2001:db8:aaaa:101::/64                        Remote  ISIS      00h03m20s  18
       2001:db8:aaaa:101::/64 (tunneled:SRV6-ISIS)                   10
2001:db8:aaaa:102::/64                        Local   SRV6      00h01m42s  3
       fe80::201-"_tmnx_fpe_2.a"                                     0
2001:db8:aaaa:102:0:1000::/128                Local   SRV6      00h01m42s  3
```

```
       Black Hole                                                    0
2001:db8:aaaa:102:0:4000::/128               Local   ISIS      00h01m42s  18
       fe80::60a:1ff:fe01:1-"int-PE-2-PE-1"                          10
2001:db8:aaaa:102:0:5000::/128               Local   ISIS      00h01m42s  18
       fe80::612:1ff:fe01:b-"int-PE-2-RR-3"                          10
-------------------------------------------------------------------------------
No. of Routes: 14
---snip---
===============================================================================
```

Verify the corresponding IPv6 BGP routes on PE-2:

```
[/]
A:admin@PE-2# show router bgp routes 2001:db8::2:5/128 hunt
===============================================================================
 BGP Router ID:2.2.2.2          AS:64500       Local AS:64500
===============================================================================
---snip---
===============================================================================
BGP IPv6 Routes
===============================================================================
-------------------------------------------------------------------------------
RIB In Entries
-------------------------------------------------------------------------------
Network        : 2001:db8::2:5/128
Nexthop        : 2001:db8::168:25:2
Path Id        : None
From           : 2001:db8::168:25:2
Res. Protocol  : LOCAL                  Res. Metric    : 0
Res. Nexthop   : 2001:db8::168:25:2
Local Pref.    : None                   Interface Name : int-PE-2-CE-2
Aggregator AS  : None                   Aggregator     : None
Atomic Aggr.   : Not Atomic             MED            : None
AIGP Metric    : None                   IGP Cost       : 0
Connector      : None
Community      : No Community Members
Cluster        : No Cluster Members
Originator Id  : None                   Peer Router Id : 2.2.2.5
Fwd Class      : None                   Priority       : None
Flags          : Used Valid Best IGP In-RTM
Route Source   : External
AS-Path        : 64505
Route Tag      : 0
Neighbor-AS    : 64505
Orig Validation: NotFound
Source Class   : 0                      Dest Class     : 0
Add Paths Send : Default
RIB Priority   : Normal
Last Modified  : 00h10m45s

-------------------------------------------------------------------------------
RIB Out Entries
-------------------------------------------------------------------------------
Network        : 2001:db8::2:5/128
Nexthop        : 2001:db8::2:2
Path Id        : None
To             : 2001:db8::2:3
Res. Protocol  : INVALID                Res. Metric    : 0
Res. Nexthop   : n/a
Local Pref.    : 100                    Interface Name : NotAvailable
Aggregator AS  : None                   Aggregator     : None
Atomic Aggr.   : Not Atomic             MED            : None
AIGP Metric    : None                   IGP Cost       : 0
```

```
Connector      : None
Community      : No Community Members
Cluster        : No Cluster Members
Originator Id  : None                  Peer Router Id : 2.2.2.3
Origin         : IGP
AS-Path        : 64505
Route Tag      : 0
Neighbor-AS    : 64505
Orig Validation: NotFound
Source Class   : 0                     Dest Class     : 0
SRv6 TLV Type  : SRv6 L3 Service TLV (5)
SRv6 SubTLV    : SRv6 SID Information (1)
Sid            : 2001:db8:aaaa:102:0:3000::
Behavior       : End.DT6 (18)
SRv6 SubSubTLV : SRv6 SID Structure (1)
Loc-Block-Len  : 48                    Loc-Node-Len   : 16
Func-Len       : 20                    Arg-Len        : 0
Tpose-Len      : 0                     Tpose-offset   : 0

-------------------------------------------------------------------------------
Routes : 2
===============================================================================
```

PE-2 then advertises route prefix 2001:db8::2:5/128, via RR-3, and inserts the SRv6 Services TLV. This TLV carries an SRv6 Service Information sub-TLV that contains the End.DT6 SID.

PE-1 receives (via RR-3) the BGP update which PE-2 sends for the IPv6 address family:

```
[/]
A:admin@PE-1# show log log-id "log_2"

---snip---
4 2022/07/19 11:39:38.404 CEST MINOR: DEBUG #2001 Base Peer 1: 2001:db8::2:3
"Peer 1: 2001:db8::2:3: UPDATE
Peer 1: 2001:db8::2:3 - Received BGP UPDATE:
    Withdrawn Length = 0
    Total Path Attr Length = 116
    Flag: 0x90 Type: 14 Len: 38 Multiprotocol Reachable NLRI:
        Address Family IPV6
        NextHop len 16 Global NextHop 2001:db8::2:2
        2001:db8::2:5/128
    Flag: 0x40 Type: 1 Len: 1 Origin: 0
    Flag: 0x40 Type: 2 Len: 6 AS Path:
        Type: 2 Len: 1 < 64505 >
    Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
    Flag: 0x80 Type: 9 Len: 4 Originator ID: 2.2.2.2
    Flag: 0x80 Type: 10 Len: 4 Cluster ID:
        3.3.3.3
    Flag: 0xc0 Type: 40 Len: 37 Prefix-SID-attr:
       SRv6 Services TLV (37 bytes):-
           Type: SRV6 L3 Service TLV (5)
           Length: 34 bytes, Reserved: 0x0
         SRv6 Service Information Sub-TLV (33 bytes)
            Type: 1 Len: 30 Rsvd1: 0x0
            SRv6 SID: 2001:db8:aaaa:102:0:3000::
            SID Flags: 0x0 Endpoint Behavior: 0x12 Rsvd2: 0x0
            SRv6 SID Sub-Sub-TLV
               Type: 1 Len: 6
               BL:48 NL:16 FL:20 AL:0 TL:0 TO:0
"
---snip---
```

Upon receipt of the BGP update from RR-3 on behalf of PE-2, PE-1 programs route prefix 2001:db8::2:5/128 in its route table and FIB. The presence of the SRv6 Services TLV indicates that the next hop is the SRv6 End.DT6 SID which, in turn, is resolved to the remote locator for PE-2.

PE-1 then advertises route prefix 2001:db8::2:5/128 to CE-1 (in RIB Out Entries).

Verify the resulting IPv6 route table on PE-1. The IPv6 route table has a route to the remote IPv6 system address of CE-2, now resolved to the End.DT6 SID of the remotely configured locator prefix of PE-2. The local auto-allocated SRv6 End.X functions have a renumbered SID, because their initial SID is now used for the statically configured SRv6 End.DT4 and SRv6 End.DT6 functions.

```
[/]
A:admin@PE-1# show router route-table ipv6

===============================================================================
IPv6 Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
---snip---
2001:db8::2:5/128                             Remote  BGP       00h00m46s  170
       2001:db8:aaaa:102:0:3000:: (tunneled:SRV6)                    10
---snip---
2001:db8:aaaa:101::/64                        Local   SRV6      00h02m02s  3
       fe80::201-"_tmnx_fpe_2.a"                                     0
2001:db8:aaaa:101:0:1000::/128                Local   SRV6      00h02m02s  3
       Black Hole                                                    0
2001:db8:aaaa:101:0:4000::/128                Local   ISIS      00h02m02s  18
       fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                          10
2001:db8:aaaa:101:0:5000::/128                Local   ISIS      00h02m02s  18
       fe80::612:1ff:fe01:1-"int-PE-1-RR-3"                          10
2001:db8:aaaa:102::/64                        Remote  ISIS      00h03m18s  18
       2001:db8:aaaa:102::/64 (tunneled:SRV6-ISIS)                   10
-------------------------------------------------------------------------------
No. of Routes: 14
---snip---
===============================================================================
```

Verify the corresponding IPv6 BGP routes on PE-1:

```
[/]
A:admin@PE-1# show router bgp routes 2001:db8::2:5/128 hunt
===============================================================================
 BGP Router ID:2.2.2.1          AS:64500       Local AS:64500
===============================================================================
---snip---
===============================================================================
BGP IPv6 Routes
===============================================================================
-------------------------------------------------------------------------------
RIB In Entries
-------------------------------------------------------------------------------
Network        : 2001:db8::2:5/128
Nexthop        : 2001:db8::2:2
Path Id        : None
From           : 2001:db8::2:3
Res. Protocol  : ISIS                  Res. Metric    : 10
Res. Nexthop   : fe80::60e:1ff:fe01:1
Local Pref.    : 100                   Interface Name : int-PE-1-PE-2
Aggregator AS  : None                  Aggregator     : None
Atomic Aggr.   : Not Atomic            MED            : None
```

```
AIGP Metric    : None                   IGP Cost       : 10
Connector      : None
Community      : No Community Members
Cluster        : 3.3.3.3
Originator Id  : 2.2.2.2                Peer Router Id : 2.2.2.3
Fwd Class      : None                   Priority       : None
Flags          : Used Valid Best IGP In-RTM
Route Source   : Internal
AS-Path        : 64505
Route Tag      : 0
Neighbor-AS    : 64505
Orig Validation: NotFound
Source Class   : 0                      Dest Class     : 0
Add Paths Send : Default
RIB Priority   : Normal
Last Modified  : 00h00m46s
SRv6 TLV Type  : SRv6 L3 Service TLV (5)
SRv6 SubTLV    : SRv6 SID Information (1)
Sid            : 2001:db8:aaaa:102:0:3000::
Behavior       : End.DT6 (18)
SRv6 SubSubTLV : SRv6 SID Structure (1)
Loc-Block-Len  : 48                     Loc-Node-Len   : 16
Func-Len       : 20                     Arg-Len        : 0
Tpose-Len      : 0                      Tpose-offset   : 0

-------------------------------------------------------------------------------
RIB Out Entries
-------------------------------------------------------------------------------
Network        : 2001:db8::2:5/128
Nexthop        : 2001:db8::168:14:1
Path Id        : None
To             : 2001:db8::168:14:2
Res. Protocol  : INVALID                Res. Metric    : 0
Res. Nexthop   : n/a
Local Pref.    : n/a                    Interface Name : NotAvailable
Aggregator AS  : None                   Aggregator     : None
Atomic Aggr.   : Not Atomic             MED            : None
AIGP Metric    : None                   IGP Cost       : 10
Connector      : None
Community      : No Community Members
Cluster        : No Cluster Members
Originator Id  : None                   Peer Router Id : 2.2.2.4
Origin         : IGP
AS-Path        : 64500 64505
Route Tag      : 0
Neighbor-AS    : 64500
Orig Validation: NotFound
Source Class   : 0                      Dest Class     : 0

-------------------------------------------------------------------------------
Routes : 2
===============================================================================
```

CE-1 receives the BGP update which PE-1 sends for the IPv6 address family:

```
[/]
A:admin@CE-1# show log log-id "log_2"

---snip---
2 2022/07/19 11:39:39.220 CEST MINOR: DEBUG #2001 Base Peer 1: 2001:db8::168:14:1
"Peer 1: 2001:db8::168:14:1: UPDATE
Peer 1: 2001:db8::168:14:1 - Received BGP UPDATE:
    Withdrawn Length = 0
```

```
    Total Path Attr Length = 59
    Flag: 0x90 Type: 14 Len: 38 Multiprotocol Reachable NLRI:
        Address Family IPV6
        NextHop len 16 Global NextHop 2001:db8::168:14:1
        2001:db8::2:5/128
    Flag: 0x40 Type: 1 Len: 1 Origin: 0
    Flag: 0x40 Type: 2 Len: 10 AS Path:
        Type: 2 Len: 2 < 64500 64505 >
"
---snip---
```

Upon receipt of the BGP update from PE-1, CE-1 programs prefix 2001:db8::2:5/128 in its route table and FIB, with the interface towards PE-1 as next hop.

Verify the resulting IPv6 route table on CE-1:

```
[/]
A:admin@CE-1# show router route-table ipv6

===============================================================================
IPv6 Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
2001:db8::2:4/128                             Local   Local     00h15m40s  0
       system                                                        0
2001:db8::2:5/128                             Remote  BGP       00h00m51s  170
       2001:db8::168:14:1                                            0
2001:db8::168:14:0/126                        Local   Local     00h15m39s  0
       int-CE-1-PE-1                                                 0
-------------------------------------------------------------------------------
No. of Routes: 3
---snip---
===============================================================================
```

Verify the corresponding IPv6 BGP routes on CE-1:

```
[/]
A:admin@CE-1# show router bgp routes 2001:db8::2:5/128 hunt
===============================================================================
 BGP Router ID:2.2.2.4          AS:64504       Local AS:64504
===============================================================================
---snip---
===============================================================================
BGP IPv6 Routes
===============================================================================
-------------------------------------------------------------------------------
RIB In Entries
-------------------------------------------------------------------------------
Network        : 2001:db8::2:5/128
Nexthop        : 2001:db8::168:14:1
Path Id        : None
From           : 2001:db8::168:14:1
Res. Protocol  : LOCAL                  Res. Metric    : 0
Res. Nexthop   : 2001:db8::168:14:1
Local Pref.    : None                   Interface Name : int-CE-1-PE-1
Aggregator AS  : None                   Aggregator     : None
Atomic Aggr.   : Not Atomic             MED            : None
AIGP Metric    : None                   IGP Cost       : 0
Connector      : None
Community      : No Community Members
Cluster        : No Cluster Members
```

```
Originator Id  : None                   Peer Router Id : 2.2.2.1
Fwd Class      : None                   Priority       : None
Flags          : Used Valid Best IGP In-RTM
Route Source   : External
AS-Path        : 64500 64505
Route Tag      : 0
Neighbor-AS    : 64500
Orig Validation: NotFound
Source Class   : 0                      Dest Class     : 0
Add Paths Send : Default
RIB Priority   : Normal
Last Modified  : 00h00m51s

-------------------------------------------------------------------------------
RIB Out Entries
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
Routes : 1
===============================================================================
```

Verify the SRv6 local SIDs on PE-2 and similar on PE-1. The SRv6 local SIDs 2001:db8:aaaa:102:0:2000:: and 2001:db8:aaaa:102:0:3000:: now correspond with the additional SRv6 End.DT4 and SRv6 End.DT6 behavior that is configured on the locator for the data transport between CE-1 and CE-2. RR-3, CE-1, and CE-2 do not have SRv6 configuration and do not have SRv6 local SIDs.

```
[/]
A:admin@PE-2# show router segment-routing-v6 local-sid

===============================================================================
Segment Routing v6 Local SIDs
===============================================================================
SID                                             Type           Function
  Locator
  Context
-------------------------------------------------------------------------------
2001:db8:aaaa:102:0:1000::                      End            1
  PE-2_loc
  Base
2001:db8:aaaa:102:0:2000::                      End.DT4        2
  PE-2_loc
  Base
2001:db8:aaaa:102:0:3000::                      End.DT6        3
  PE-2_loc
  Base
2001:db8:aaaa:102:0:4000::                      End.X          4
  PE-2_loc
    None
2001:db8:aaaa:102:0:5000::                      End.X          5
  PE-2_loc
    None
-------------------------------------------------------------------------------
SIDs : 5
-------------------------------------------------------------------------------
===============================================================================
```

Verify the SRv6 base routing instance on PE-2 and similar on PE-1.

```
[/]
A:admin@PE-2# show router segment-routing-v6 base-routing-instance

===============================================================================
Segment Routing v6 Base Routing Instance
```

```
===============================================================================
Locator
  Type       Function    SID                                      Status/InstId
    SRH-mode Protection  Interface
-------------------------------------------------------------------------------
PE-2_loc
  End.DT4              2 2001:db8:aaaa:102:0:2000::              ok
  End.DT6              3 2001:db8:aaaa:102:0:3000::              ok
  End                  1 2001:db8:aaaa:102:0:1000::              ok
    USP
-------------------------------------------------------------------------------
  Auto-allocated End.X: PSP Unprotected,
-------------------------------------------------------------------------------
  End.X               *4 2001:db8:aaaa:102:0:4000::              0
    PSP      Unprotected int-PE-2-PE-1
    ISIS Level: L2 Mac Address: 04:0a:01:01:00:01 Nbr Sys Id: 0010.0100.1001
  End.X               *5 2001:db8:aaaa:102:0:5000::              0
    PSP      Unprotected int-PE-2-RR-3
    ISIS Level: L2 Mac Address: 04:12:01:01:00:0b Nbr Sys Id: 0010.0100.1003
-------------------------------------------------------------------------------
===============================================================================
Legend: * - System allocated
```

Verify that the tunnel from PE-1 to the remote locator has SRv6 encapsulation and similar for the tunnel from PE-2 to the remote locator. The tunnel tables on RR-3 and on CE-1 are empty.

```
[/]
A:admin@PE-1# show router tunnel-table ipv6

===============================================================================
IPv6 Tunnel Table (Router: Base)
===============================================================================
Destination                                 Owner     Encap TunnelId  Pref
Nexthop                                     Color           Metric
-------------------------------------------------------------------------------
2001:db8:aaaa:102::/64                      srv6-isis SRV6  524289    0
  fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                        10
-------------------------------------------------------------------------------
---snip---
===============================================================================
```

Verify that the tunnel from PE-1 to the remote locator uses the “int-PE-1-PE-2” interface and similar for the tunnel from PE-2 to the remote locator, where that tunnel uses the “int-PE-2-PE-1” interface. Interface “int-PE-1-PE-2” is configured on port 1/1/c1/1:1000. The FP tunnel tables on RR-3 and on CE-1 are empty.

```
[/]
A:admin@PE-1# show router fp-tunnel-table 1 ipv6

===============================================================================
IPv6 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                 Protocol         Tunnel-ID
  Lbl/SID
    NextHop                                                  Intf/Tunnel
  Lbl/SID (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
2001:db8:aaaa:102::/64                      SRV6              524289
```

```
  -
    fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                        1/1/c1/1:1000
-------------------------------------------------------------------------------
Total Entries : 1
-------------------------------------------------------------------------------
===============================================================================
```

Verify that data transport is possible between CE-1 and CE-2. IPv4 data flows from CE-1 to PE-1, where it is SRv6-encapsulated and forwarded via the SRv6 tunnel to PE-2. At PE-2, the data is decapsulated and is forwarded to CE-2. Between PE-1 and PE-2, the IPv4 data cannot flow unencapsulated because there is no IPv4 interface between PE-1 and PE-2.

```
[/]
A:admin@CE-1# ping 192.0.2.5
PING 192.0.2.5 56 data bytes
64 bytes from 192.0.2.5: icmp_seq=1 ttl=62 time=2.13ms.
---snip---
---- 192.0.2.5 PING Statistics ----
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min = 1.76ms, avg = 1.98ms, max = 2.35ms, stddev = 0.225ms
```

```
[/]
A:admin@CE-1# traceroute 192.0.2.5
traceroute to 192.0.2.5, 30 hops max, 40 byte packets
  1  172.16.14.1 (172.16.14.1)    0.783 ms  0.864 ms  0.893 ms
  2  0.0.0.0  * * *
  3  192.0.2.5 (192.0.2.5)    2.07 ms  1.83 ms  1.87 ms
```

IPv6 data flows from CE-1 to PE-1, where it is SRv6 encapsulated and forwarded via the SRv6 tunnel to PE-2. At PE-2, the data is decapsulated and is forwarded to CE-2. The IPv6 data does not flow with native IPv6 between PE-1 and PE-2 because then it would use the 2001:db8::168:12:1 IPv6 interface instead of the 2001:db8::2:2 IPv6 system address in the second hop. The same is true for data transport between CE-2 and CE-1.

```
[/]
A:admin@CE-1# ping 2001:db8::2:5
PING 2001:db8::2:5 56 data bytes
64 bytes from 2001:db8::2:5 icmp_seq=1 hlim=62 time=1.68ms.
---snip---
---- 2001:db8::2:5 PING Statistics ----
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min = 1.68ms, avg = 1.75ms, max = 1.83ms, stddev = 0.056ms
```

```
[/]
A:admin@CE-1# traceroute 2001:db8::2:5
traceroute to 2001:db8::2:5, 30 hops max, 60 byte packets
  1  2001:db8::168:14:1 (2001:db8::168:14:1)    0.835 ms  0.915 ms  0.809 ms
  2  2001:db8::2:2 (2001:db8::2:2)    1.54 ms  1.59 ms  1.51 ms
  3  2001:db8::2:5 (2001:db8::2:5)    1.88 ms  1.73 ms  2.00 ms
```

Conclusion

SRv6 Encapsulation in the base routing instance can be used to transport native IPv4 and IPv6 data across an SRv6-enabled provider network.

SRv6 Loop-Free Alternate

This chapter provides information about loop-free alternate for segment routing over IPv6 .

Topics in this chapter include:

Applicability

The information and configuration in this chapter are based on SR OS Release 22.2.R1. Segment routing over IPv6 (SRv6) is supported on FP4-based equipment in SR OS Release 21.5.R2 and later.

Overview

SR OS Release 21.5.R2 and later support loop-free alternate (LFA) for segment routing over IPv6 (SRv6). This includes regular LFA, remote LFA (R-LFA) and topology independent LFA (TI-LFA) for routers in a service originating role and for routers in a transit role, with or without segment termination.

The local router installs its locator prefix in its IPv6 route table and IPv6 forwarding information base (FIB), and advertises its locator prefix in IS-IS with the SRv6 locator sub-TLV. Each remote router populates its IPv6 route table and IPv6 FIB with the received locator prefixes, including the tunneled next hop to the originating router.

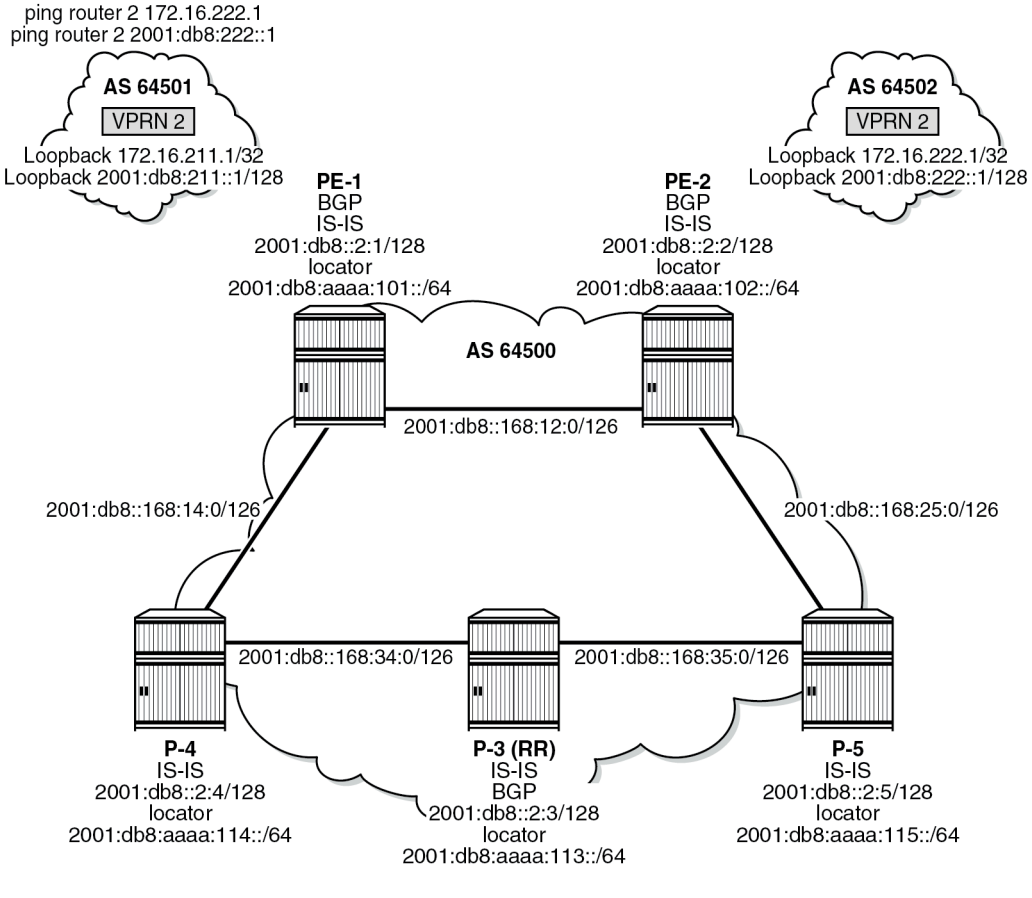
The LFA backup path for a local End.X segment identifier (SID) or a local LAN End.X SID is programmed in the IPv6 route table and in the IPv6 FIB with the specific entry corresponding to the local locator prefix.

The LFA backup path for a remote locator prefix entry is programmed in the IPv6 route table and in the IPv6 FIB. The LFA backup path for a remote End SID, End.DT4 SID, End.DT6 SID, or End.DX2 SID uses that remote locator prefix.

Configuration

Figure 65: Example topology shows the example topology with five SRv6-capable routers. The SRv6-enabled network that it represents comprises PE-1, PE-2, and P-3 in the control and data planes, and P-4 and P-5 in the data plane only. The SRv6-enabled network has only IPv6 addresses and interfaces.

Figure 65: Example topology



For the transport of IPv4 and IPv6 data from the VPRN on PE-1 to the VPRN on PE-2, PE-1 acts as the SRv6 ingress PE node, while PE-2 acts as the SRv6 egress PE node.

SRv6 and forwarding path extension (FPE) are configured on all routers. P-3 acts as the BGP route reflector in the control plane. As long as the link between PE-1 and PE-2 is operational, P-3 does not participate in the SRv6 data transport between PE-1 and PE-2. When the link between PE-1 and PE-2 fails, SRv6 data transport uses an LFA backup path via P-3.

The **ping** and **traceroute** commands between IPv4 and IPv6 loopback addresses in the VPRNs simulate data transport.

SRv6 for VPRN is established between PE-1 and PE-2, as described in the Segment Routing over IPv6 for VPRN chapter. The metric on all links, except one, is set to 10. When the metric on the link between PE-2 and P-5 is set to 21, configuring TI-LFA on PE-1 for all destination prefixes using the protected link PE-1–PE-2, results in a PQ-router P-5. In this case, the End SID of P-5 suffices. When the metric on the link between P-3 and P-5 is set to 21, configuring TI-LFA on PE-1 for all destination prefixes using the protected link PE-1-PE-2, results in disjointed P-router P-3 and Q-router P-5. In that case, the End.X SID referencing the interface on P-3 facing P-5 suffices to reach the Q node.

Configure the router

This configuration includes:

- ports and IPv6-only interfaces on all routers
- port cross-connect (PXC) on all routers, using internal loopbacks on an FP4 MAC chip, as described in the Segment Routing over IPv6 chapter
- IS-IS on all routers, including:
  - level 2 capability with wide metrics (for the 128-bit identifiers)
  - level 2 metric is 10 on all IS-IS interfaces, but 21 on the IS-IS interface between PE-2 and P-5
  - native IPv6 routing
  - as a best practice to advertise the router capability within the autonomous system (AS), also configure:
    - **traffic-engineering**
    - **traffic-engineering-options**
- BGP on PE-1, PE-2, and P-3, with internal group “gr_v6_internal” that includes:
  - IPv4, IPv6, VPN-IPv4 and VPN-IPv6 families
  - **extended-nh-encoding** for IPv4 and VPN-IPv4
  - **advertise-ipv6-next-hops** for IPv4, VPN-IPv4 and VPN-IPv6
  - BGP neighbor **system** IPv6 addresses
  - On PE-1 and PE-2 only: **next-hop-self**

The core network topology uses IPv6 for BGP peering (with 16 byte next hop addresses), so to advertise and receive IPv4 routes (which have 4 byte next hop addresses) with IPv6 next hop addresses, the commands **advertise-ipv6-next-hops** and **extended-nh-encoding** need to be configured at the BGP, group, or neighbor level. The **advertise-ipv6-next-hops** command instructs the system to advertise IPv4 routes with IPv6 next hop addresses. The **extended-nh-encoding** command configures BGP to advertise the capability to receive IPv4 routes with IPv6 next hop addresses.

The following example configuration applies for PE-1 and is similar for the other routers, with the following differences:

- P-3 acts as a BGP route reflector
- BGP is not configured on P-4 and P-5

```
[/]
A:admin@PE-1# configure {
    router "Base" {
        autonomous-system 64500
        interface "int-PE-1-P-4" {
            description "interface between PE-1 and P-4"
            port 1/1/c2/1:1000
            ipv6 {
                address 2001:db8::168:14:1 {
                    prefix-length 126
                }
            }
        }
        interface "int-PE-1-PE-2" {
```

```
            description "interface between PE-1 and PE-2"
            port 1/1/c1/1:1000
            ipv6 {
                address 2001:db8::168:12:1 {
                    prefix-length 126
                }
            }
        }
        interface "system" {
            description "system interface of PE-1"
            ipv6 {
                address 2001:db8::2:1 {
                    prefix-length 128
                }
            }
        }
        isis 0 {
            admin-state enable
            advertise-router-capability as
            ipv6-routing native
            level-capability 2    # required for SRv6
            router-id 1.1.1.1    # must be unique and in the format of an IPv4 address
            traffic-engineering true
            area-address [49.0001]
            traffic-engineering-options {
                ipv6 true
                application-link-attributes {
                }
            }
            interface "int-PE-1-P-4" {
                interface-type point-to-point
                level 1 {
                    metric 10
                }
                level 2 {
                    metric 10
                }
            }
            interface "int-PE-1-PE-2" {
                interface-type point-to-point
                level 1 {
                    metric 10
                }
                level 2 {
                    metric 10
                }
            }
            interface "system" {
                passive true
            }
            level 2 {
                wide-metrics-only true    # required for SRv6
            }
        }
        bgp {
            min-route-advertisement 1
            router-id 2.2.2.1    # must be unique and in the format of an IPv4 address
            rapid-withdrawal true
            split-horizon true
            ebgp-default-reject-policy {
                import false
                export false
            }
            rapid-update {
```

```
                vpn-ipv4 true
                vpn-ipv6 true
            }
            group "gr_v6_internal" {
                description "internal bgp group on PE-1"
                next-hop-self true
                type internal
                family {
                    ipv4 true
                    vpn-ipv4 true
                    ipv6 true
                    vpn-ipv6 true
                }
                extended-nh-encoding {
                    vpn-ipv4 true
                    ipv4 true
                }
                advertise-ipv6-next-hops {
                    vpn-ipv6 true
                    vpn-ipv4 true
                    ipv4 true
                }
            }
            neighbor "2001:db8::2:3" {    # P-3 system address
                group "gr_v6_internal"
            }
        }
        exit all
```

Note:
Do not advertise tunnel links, because that enables forwarding adjacencies. IS-IS does not compute a remote LFA or a TI-LFA backup for an SR-ISIS tunnel when forwarding adjacency (configured via the **advertise-tunnel-links** command) is enabled in the IS-IS instance, even if these two types of LFAs are enabled in the configuration of that same IS-IS instance.

Configure the VPRN services on PE-1 and on PE-2

This configuration includes:

- an IPv4 address and an IPv6 address for a loopback interface "lb_itf_vprn"
- BGP, with external group “gr_v6_vprn” that includes the following capabilities:
  - IPv4 and IPv6 families
  - **extended-nh-encoding** for IPv4
  - **advertise-ipv6-next-hops** for IPv4
  - BGP neighbor **interface** IPv6 addresses, with BGP neighbors in a different external AS

The following example configuration applies for VPRN 2 on PE-1 and is similar for VPRN 2 on PE-2.

```
[/]
A:admin@PE-1# configure {
    service {
        vprn "VPRN_2" {
            admin-state enable
            description "VPRN 2 on PE-1"
            service-id 2
            customer "1"
```

```
            autonomous-system 64500
            bgp {
                ebgp-default-reject-policy {
                    import false
                    export false
                }
                group "gr_v6_vprn" {
                    description "external bgp group for VPRN 2 on PE-1"
                    family {
                        ipv4 true
                        ipv6 true
                    }
                    extended-nh-encoding {
                        ipv4 true
                    }
                    advertise-ipv6-next-hops {
                        ipv4 true
                    }
                }
                neighbor "2001:db8:101::1" {
                    group "gr_v6_vprn"
                    type external
                    peer-as 64501
                }
            }
            interface "lb_itf_vprn" {
                description "VPRN 2 interface on PE-1 for external subnet"
                loopback true
                ipv4 {
                    primary {
                        address 172.16.211.1
                        prefix-length 32
                    }
                }
                ipv6 {
                    address 2001:db8:211::1 {
                        prefix-length 128
                    }
                }
            }
            exit all
```

Configure SRv6 in the router Base context on all routers

Configure the locator in the **router Base segment-routing segment-routing-v6** context on PE-2 and similar on the other routers, with different **ip-prefix** for the locators.

```
[/]
A:admin@PE-2# configure {
    router "Base" {
        segment-routing {
            segment-routing-v6 {
                locator "PE-2_loc" {
                    admin-state enable
                    block-length 48
                    prefix {
                        ip-prefix 2001:db8:aaaa:102::/64
                    }
                }
            }
```

```
            exit all
```

Configure the FPEs on PE-2 and identical on the other routers.

```
[/]
A:admin@PE-2# configure {
    fwd-path-ext {
        fpe 1 {
            path {
                pxc 1
            }
            application {
                srv6 {
                    type origination
                }
            }
        }
        fpe 2 {
            path {
                pxc 2
            }
            application {
                srv6 {
                    type termination
                }
            }
        }
        exit all
```

Use FPE 1 as the SRv6 origination FPE in the **router Base segment-routing segment-routing-v6** context and FPE 2 as the SRv6 termination FPE in the **router Base segment-routing segment-routing-v6 locator** context on PE-2. The configuration is similar on the other routers, with different locators. For more information, see the Segment Routing over IPv6 chapter.

```
[/]
A:admin@PE-2# configure {
    router "Base" {
        segment-routing {
            segment-routing-v6 {
                origination-fpe [1]
                locator "PE-2_loc" {
                    admin-state enable
                    termination-fpe [2]
                }
            }
            exit all
```

Configure the SRv6 End function (equivalent to an IPv4 node SID) and SRv6 End.X functions (equivalent to IPv4 adjacency SIDs) in the **router Base segment-routing segment-routing-v6 base-routing-instance locator** context on all routers, with different locators.

```
[/]
A:admin@PE-2# configure {
    router "Base" {
        mpls-labels {
            sr-labels {
                start 20000
                end 20999
            }
            reserved-label-block "SRv6" {
                start-label 30100
```

```
                end-label 30199
            }
        }
        segment-routing {
            segment-routing-v6 {
                locator "PE-2_loc" {
                    static-function {
                        max-entries 3
                        label-block "SRv6"
                    }
                }
                base-routing-instance {
                    locator "PE-2_loc" {
                        function {
                            end 1 {
                                srh-mode usp
                            }
                            end-x-auto-allocate usp protection protected { }
                        }
                    }
                }
            }
        }
        exit all
```

While not strictly needed, allow for three static functions. New SRv6 functions (for example End.DT4 and End.DT6), can then be configured without needing to reshuffle the automatic SRv6 function numbering. Ensure that the End.X functions have protection on. As a result, the End.X functions are only instantiated when **loopfree-alternate** is configured in the **router Base isis** context.

Advertise the locator in IS-IS while ensuring level 2 capability on PE-2. Configure other routers similarly, with different locators.

```
[/]
A:admin@PE-2# configure {
    router "Base" {
        isis 0 {
            segment-routing-v6 {
                admin-state enable
                locator "PE-2_loc" {
                    level-capability 2
                }
            }
            exit all
```

Use the **show router segment-routing-v6 summary** command to verify the locator and origination FPE configuration.

Configure SRv6 for the VPRNs on PE-1 and on PE-2

Create an SRv6 instance for the VPRN service. Use the locator from the **router Base segment-routing segment-routing-v6** context and configure End.DT4 and End.DT6 functions for it.

Use the created SRv6 instance in the **service vprn bgp-ipvpn segment-routing-v6** context, with the configured locator as the default locator. Ensure a unique route distinguisher. Use the unique PE-2 system

IPv6 address as the source address. Use a similar configuration on PE-1, with the PE-1 locator as default locator, the PE-1 system IPv6 address as the source address, and a different route distinguisher.

```
[/]
A:admin@PE-2# configure {
    service {
        vprn "VPRN_2" {
            segment-routing-v6 1 {
                locator "PE-2_loc" {
                    function {
                        end-dt4 {
                        }
                        end-dt6 {
                        }
                    }
                }
            }
            bgp-ipvpn {
                segment-routing-v6 1 {
                    admin-state enable
                    route-distinguisher "192.0.2.2:2"
                    source-address 2001:db8::2:2
                    vrf-target {
                        community "target:64506:2"
                    }
                    srv6 {
                        instance 1
                        default-locator "PE-2_loc"
                    }
                }
            }
            exit all
```

This configuration results in BGP update exchanges between PE-2 and PE-1, via P-3, and between PE-1 and PE-2, via P-3.

At this point, verify that data transport is possible between the local VPRN on PE-1 and the remote VPRN on PE-2.

```
[/]
A:admin@PE-1# ping 172.16.222.1 router-instance "VPRN_2"
PING 172.16.222.1 56 data bytes
64 bytes from 172.16.222.1: icmp_seq=1 ttl=64 time=1.85ms.
---snip---
---- 172.16.222.1 PING Statistics ----
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min = 1.74ms, avg = 1.82ms, max = 1.93ms, stddev = 0.065ms
```

```
[/]
A:admin@PE-1# traceroute 172.16.222.1 router-instance "VPRN_2"
traceroute to 172.16.222.1, 30 hops max, 40 byte packets
  1  172.16.222.1 (172.16.222.1)    2.18 ms  1.71 ms  1.76 ms
```

```
[/]
A:admin@PE-1# ping 2001:db8:222::1 router-instance "VPRN_2"
PING 2001:db8:222::1 56 data bytes
64 bytes from 2001:db8:222::1 icmp_seq=1 hlim=64 time=1.54ms.
---snip---
---- 2001:db8:222::1 PING Statistics ----
5 packets transmitted, 5 packets received, 0.00% packet loss
```

```
round-trip min = 1.54ms, avg = 1.60ms, max = 1.73ms, stddev = 0.072ms
```

```
[/]
A:admin@PE-1# traceroute 2001:db8:222::1 router-instance "VPRN_2"
traceroute to 2001:db8:222::1, 30 hops max, 60 byte packets
  1  2001:db8:222::1 (2001:db8:222::1)    1.83 ms  1.80 ms  1.74 ms
```

The result of the verification complies with the route tables for the local VPRN on PE-1, which contains routes for the loopback addresses in the remote VPRN on PE-2. The same is true for data transport between the remote VPRN on PE-2 and the local VPRN on PE-1.

```
[/]
A:admin@PE-1# show router 2 route-table ipv4

===============================================================================
Route Table (Service: 2)
===============================================================================
Dest Prefix[Flags]                              Type    Proto     Age        Pref
      Next Hop[Interface Name]                                      Metric
-------------------------------------------------------------------------------
172.16.211.1/32                                 Local   Local     00h21m39s  0
       lb_itf_vprn                                                    0
172.16.222.1/32                                 Remote  BGP VPN   00h02m28s  170
       2001:db8:aaaa:102:78a6:c000:: (tunneled:SRV6)                  10
-------------------------------------------------------------------------------
No. of Routes: 2
---snip---
===============================================================================
```

```
[/]
A:admin@PE-1# show router 2 route-table ipv6

===============================================================================
IPv6 Route Table (Service: 2)
===============================================================================
Dest Prefix[Flags]                              Type    Proto     Age        Pref
      Next Hop[Interface Name]                                      Metric
-------------------------------------------------------------------------------
2001:db8:211::1/128                             Local   Local     00h21m37s  0
       lb_itf_vprn                                                    0
2001:db8:222::1/128                             Remote  BGP VPN   00h02m28s  170
       2001:db8:aaaa:102:78a6:b000:: (tunneled:SRV6)                  10
-------------------------------------------------------------------------------
No. of Routes: 2
---snip---
===============================================================================
```

The IPv4 route table and IPv4 FIB remain empty, while the IPv6 route table and IPv6 FIB contain routes for the local, IS-IS, and SRv6 protocols. The remote destinations to PE-2 and P-5 are reached via the “int-PE-1-PE-2” interface. There are no routes yet for the local End.X functions. The local End.X functions are not yet instantiated, because there is no regular LFA protection while protection is enabled for End.X function. Verify the IPv6 route table.

```
[/]
A:admin@PE-1# show router route-table ipv6

===============================================================================
IPv6 Route Table (Router: Base)
===============================================================================
```

```
Dest Prefix[Flags]                              Type    Proto     Age        Pref
      Next Hop[Interface Name]                                      Metric
-------------------------------------------------------------------------------
2001:db8::2:1/128                               Local   Local     00h31m23s  0
       system                                                          0
2001:db8::2:2/128                               Remote  ISIS      00h08m15s  18
       fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                            10
2001:db8::2:3/128                               Remote  ISIS      00h08m15s  18
       fe80::616:1ff:fe01:1-"int-PE-1-P-4"                             20
2001:db8::2:4/128                               Remote  ISIS      00h08m15s  18
       fe80::616:1ff:fe01:1-"int-PE-1-P-4"                             10
2001:db8::2:5/128                               Remote  ISIS      00h08m15s  18
       fe80::616:1ff:fe01:1-"int-PE-1-P-4"                             30
2001:db8::168:12:0/126                          Local   Local     00h31m22s  0
       int-PE-1-PE-2                                                   0
2001:db8::168:14:0/126                          Local   Local     00h31m22s  0
       int-PE-1-P-4                                                    0
2001:db8::168:25:0/126                          Remote  ISIS      00h08m15s  18
       fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                            31
2001:db8::168:34:0/126                          Remote  ISIS      00h08m15s  18
       fe80::616:1ff:fe01:1-"int-PE-1-P-4"                             20
2001:db8::168:35:0/126                          Remote  ISIS      00h08m15s  18
       fe80::616:1ff:fe01:1-"int-PE-1-P-4"                             30
2001:db8:aaaa:101::/64                          Local   SRV6      00h10m40s  3
       fe80::201-"_tmnx_fpe_2.a"                                       0
2001:db8:aaaa:101:0:1000::/128                  Local   SRV6      00h10m40s  3
       Black Hole                                                      0
2001:db8:aaaa:102::/64                          Remote  ISIS      00h08m01s  18
       2001:db8:aaaa:102::/64 (tunneled:SRV6-ISIS)                     10
2001:db8:aaaa:113::/64                          Remote  ISIS      00h07m51s  18
       2001:db8:aaaa:113::/64 (tunneled:SRV6-ISIS)                     20
2001:db8:aaaa:114::/64                          Remote  ISIS      00h07m30s  18
       2001:db8:aaaa:114::/64 (tunneled:SRV6-ISIS)                     10
2001:db8:aaaa:115::/64                          Remote  ISIS      00h07m08s  18
       2001:db8:aaaa:115::/64 (tunneled:SRV6-ISIS)                     30
-------------------------------------------------------------------------------
No. of Routes: 16
---snip---
===============================================================================
```

Verify the corresponding IPv6 FIB.

```
[/]
A:admin@PE-1# show router fib 1 ipv6

===============================================================================
FIB Display
===============================================================================
Prefix [Flags]                                                  Protocol
  NextHop
-------------------------------------------------------------------------------
2001:db8::2:1/128                                               LOCAL
  2001:db8::2:1 (system)
2001:db8::2:2/128                                               ISIS
  fe80::60e:1ff:fe01:1 (int-PE-1-PE-2)
2001:db8::2:3/128                                               ISIS
  fe80::616:1ff:fe01:1 (int-PE-1-P-4)
2001:db8::2:4/128                                               ISIS
  fe80::616:1ff:fe01:1 (int-PE-1-P-4)
2001:db8::2:5/128                                               ISIS
  fe80::616:1ff:fe01:1 (int-PE-1-P-4)
2001:db8::168:12:0/126                                          LOCAL
  2001:db8::168:12:0 (int-PE-1-PE-2)
```

```
2001:db8::168:14:0/126                                  LOCAL
  2001:db8::168:14:0 (int-PE-1-P-4)
2001:db8::168:25:0/126                                  ISIS
  fe80::60e:1ff:fe01:1 (int-PE-1-PE-2)
2001:db8::168:34:0/126                                  ISIS
  fe80::616:1ff:fe01:1 (int-PE-1-P-4)
2001:db8::168:35:0/126                                  ISIS
  fe80::616:1ff:fe01:1 (int-PE-1-P-4)
2001:db8:aaaa:101::/64                                  SRV6
  fe80::201 (_tmnx_fpe_2.a)
2001:db8:aaaa:101:0:1000::/128                          SRV6
  Blackhole
2001:db8:aaaa:102::/64                                  ISIS
  2001:db8:aaaa:102::/64 (Transport:SRV6:524289)
2001:db8:aaaa:113::/64                                  ISIS
  2001:db8:aaaa:113::/64 (Transport:SRV6:524290)
2001:db8:aaaa:114::/64                                  ISIS
  2001:db8:aaaa:114::/64 (Transport:SRV6:524291)
2001:db8:aaaa:115::/64                                  ISIS
  2001:db8:aaaa:115::/64 (Transport:SRV6:524292)
-------------------------------------------------------------------------------
Total Entries : 16
-------------------------------------------------------------------------------
===============================================================================
```

The IS-IS data base contains the following information. Only the End functions are already instantiated, on their respective locators.

```
[/]
A:admin@PE-1# show router isis database detail

===============================================================================
Rtr Base ISIS Instance 0 Database (detail)
===============================================================================

Displaying Level 1 database
-------------------------------------------------------------------------------
Level (1) LSP Count : 0

Displaying Level 2 database
-------------------------------------------------------------------------------
LSP ID    : PE-1.00-00                                  Level     : L2
---snip---
SYS ID    : 0010.0100.1001         SysID Len : 6        Used Len  : 368

TLVs :
  ---snip---
  Router Cap : 1.1.1.1, D:0, S:0
    TE Node Cap : B E M  P
    SRv6 Cap: 0x0000
    SR Alg: metric based SPF
    Node MSD Cap: BMI : 0 SRH-MAX-SL : 10 SRH-MAX-END-POP : 9 SRH-MAX-H-ENCAPS : 1 SRH-MAX-END-
D : 9
  ---snip---
  TE IS Nbrs   :
    Nbr   : PE-2.00
    Default Metric  : 10
    ---snip---
  TE IS Nbrs   :
    Nbr   : P-4.00
    Default Metric  : 10
    ---snip---
  IPv6 Reach:
```

```
    ---snip---
    Metric: ( I ) 0
    Prefix   : 2001:db8:aaaa:101::/64
  SRv6 Locator  :
    MT ID : 0
    Metric: ( ) 0 Algo:0
    Prefix   : 2001:db8:aaaa:101::/64
    Sub TLV   :
      End-SID   : 2001:db8:aaaa:101:0:1000::, flags:0x0, endpoint:End-USP

-------------------------------------------------------------------------------
LSP ID    : PE-2.00-00                                  Level     : L2
---snip---
SYS ID    : 0010.0100.1002         SysID Len : 6        Used Len  : 368

TLVs :
  ---snip---
  Router Cap : 1.1.1.2, D:0, S:0
    TE Node Cap : B E M  P
    SRv6 Cap: 0x0000
    SR Alg: metric based SPF
    Node MSD Cap: BMI : 0 SRH-MAX-SL : 10 SRH-MAX-END-POP : 9 SRH-MAX-H-ENCAPS : 1 SRH-MAX-END-
D : 9
  ---snip---
  TE IS Nbrs   :
    Nbr   : PE-1.00
    Default Metric  : 10
    ---snip---
  TE IS Nbrs   :
    Nbr   : P-5.00
    Default Metric  : 21
    ---snip---
  IPv6 Reach:
    ---snip---
    Metric: ( I ) 21
    Prefix   : 2001:db8::168:25:0/126
    Metric: ( I ) 0
    Prefix   : 2001:db8:aaaa:102::/64
  SRv6 Locator  :
    MT ID : 0
    Metric: ( ) 0 Algo:0
    Prefix   : 2001:db8:aaaa:102::/64
    Sub TLV   :
      End-SID   : 2001:db8:aaaa:102:0:1000::, flags:0x0, endpoint:End-USP

-------------------------------------------------------------------------------
LSP ID    : P-3.00-00                                   Level     : L2
---snip---
SYS ID    : 0010.0100.1003         SysID Len : 6        Used Len  : 367

TLVs :
  ---snip---
  Router Cap : 1.1.1.3, D:0, S:0
    TE Node Cap : B E M  P
    SRv6 Cap: 0x0000
    SR Alg: metric based SPF
    Node MSD Cap: BMI : 0 SRH-MAX-SL : 10 SRH-MAX-END-POP : 9 SRH-MAX-H-ENCAPS : 1 SRH-MAX-END-
D : 9
  ---snip---
  TE IS Nbrs   :
    Nbr   : P-4.00
    Default Metric  : 10
    ---snip---
  TE IS Nbrs   :
```

```
    Nbr   : P-5.00
    Default Metric  : 10
    ---snip---
  IPv6 Reach:
    ---snip---
    Metric: ( I ) 0
    Prefix   : 2001:db8:aaaa:113::/64
  SRv6 Locator  :
    MT ID : 0
    Metric: ( ) 0 Algo:0
    Prefix   : 2001:db8:aaaa:113::/64
    Sub TLV   :
      End-SID   : 2001:db8:aaaa:113:0:1000::, flags:0x0, endpoint:End-USP

-------------------------------------------------------------------------------
LSP ID    : P-4.00-00                                   Level     : L2
---snip---
SYS ID    : 0010.0100.1004         SysID Len : 6        Used Len  : 367

TLVs :
  ---snip---
  Router Cap : 1.1.1.4, D:0, S:0
    TE Node Cap : B E M  P
    SRv6 Cap: 0x0000
    SR Alg: metric based SPF
    Node MSD Cap: BMI : 0 SRH-MAX-SL : 10 SRH-MAX-END-POP : 9 SRH-MAX-H-ENCAPS : 1 SRH-MAX-END-
D : 9
  ---snip---
  TE IS Nbrs   :
    Nbr   : P-3.00
    Default Metric  : 10
    ---snip---
  TE IS Nbrs   :
    Nbr   : PE-1.00
    Default Metric  : 10
    ---snip---
  IPv6 Reach:
    ---snip---
    Metric: ( I ) 0
    Prefix   : 2001:db8:aaaa:114::/64
  SRv6 Locator  :
    MT ID : 0
    Metric: ( ) 0 Algo:0
    Prefix   : 2001:db8:aaaa:114::/64
    Sub TLV   :
      End-SID   : 2001:db8:aaaa:114:0:1000::, flags:0x0, endpoint:End-USP

-------------------------------------------------------------------------------
LSP ID    : P-5.00-00                                   Level     : L2
---snip---
SYS ID    : 0010.0100.1005         SysID Len : 6        Used Len  : 367

TLVs :
  ---snip---
  Router Cap : 1.1.1.5, D:0, S:0
    TE Node Cap : B E M  P
    SRv6 Cap: 0x0000
    SR Alg: metric based SPF
    Node MSD Cap: BMI : 0 SRH-MAX-SL : 10 SRH-MAX-END-POP : 9 SRH-MAX-H-ENCAPS : 1 SRH-MAX-END-
D : 9
  ---snip---
  TE IS Nbrs   :
    Nbr   : P-3.00
    Default Metric  : 10
```

```
    ---snip---
  TE IS Nbrs   :
    Nbr   : PE-2.00
    Default Metric  : 21
    ---snip---
  IPv6 Reach:
    ---snip---
    Metric: ( I ) 21
    Prefix   : 2001:db8::168:25:0/126
    ---snip---
    Metric: ( I ) 0
    Prefix   : 2001:db8:aaaa:115::/64
  SRv6 Locator  :
    MT ID : 0
    Metric: ( ) 0 Algo:0
    Prefix   : 2001:db8:aaaa:115::/64
    Sub TLV   :
      End-SID   : 2001:db8:aaaa:115:0:1000::, flags:0x0, endpoint:End-USP

Level (2) LSP Count : 5
-------------------------------------------------------------------------------
---snip---
===============================================================================
```

Verify the SRv6 local SIDs and SRv6 base routing instances on PE-1 and similar on PE-2. The End.X functions are not yet instantiated.

```
[/]
A:admin@PE-1# show router segment-routing-v6 local-sid

===============================================================================
Segment Routing v6 Local SIDs
===============================================================================
SID                                                 Type           Function
  Locator
  Context
-------------------------------------------------------------------------------
2001:db8:aaaa:101:0:1000::                          End            1
  PE-1_loc
  Base
2001:db8:aaaa:101:78a6:b000::                       End.DT6        494187
  PE-1_loc
  SvcId: 2 Name: VPRN_2
2001:db8:aaaa:101:78a6:c000::                       End.DT4        494188
  PE-1_loc
  SvcId: 2 Name: VPRN_2
-------------------------------------------------------------------------------
SIDs : 3
-------------------------------------------------------------------------------
===============================================================================
```

The End.X functions not yet being instantiated can also be verified in the SRv6 base routing instances on PE-1 and similar on PE-2. Only the End function is already instantiated.

```
[/]
A:admin@PE-1# show router segment-routing-v6 base-routing-instance

===============================================================================
Segment Routing v6 Base Routing Instance
===============================================================================
Locator
  Type       Function    SID                                     Status/InstId
```

```
    SRH-mode Protection  Interface
-------------------------------------------------------------------------------
PE-1_loc
  End                 1 2001:db8:aaaa:101:0:1000::             ok
    USP
-------------------------------------------------------------------------------
  Auto-allocated End.X: USP Protected,
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
===============================================================================
Legend: * - System allocated
```

Verify that the tunnels have SRv6 encapsulation.

```
[/]
A:admin@PE-1# show router tunnel-table ipv6

===============================================================================
IPv6 Tunnel Table (Router: Base)
===============================================================================
Destination                                 Owner     Encap TunnelId  Pref
Nexthop                                     Color           Metric
-------------------------------------------------------------------------------
2001:db8:aaaa:102::/64                      srv6-isis SRV6  524289    0
  fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                        10
2001:db8:aaaa:113::/64                      srv6-isis SRV6  524290    0
  fe80::616:1ff:fe01:1-"int-PE-1-P-4"                         20
2001:db8:aaaa:114::/64                      srv6-isis SRV6  524291    0
  fe80::616:1ff:fe01:1-"int-PE-1-P-4"                         10
2001:db8:aaaa:115::/64                      srv6-isis SRV6  524292    0
  fe80::616:1ff:fe01:1-"int-PE-1-P-4"                         30
-------------------------------------------------------------------------------
---snip---
===============================================================================
```

Verify the interfaces that the tunnels are using. Interface “int-PE-1-PE-2” is configured on port 1/1/c1/1:1000. Interface “int-PE-1-P-4” is configured on port 1/1/c2/1:1000. SRv6 data is transported to PE-2 over the link between PE-1 and PE-2, via next hop fe80::60e:1ff:fe01:1-"int-PE-1-PE-2".

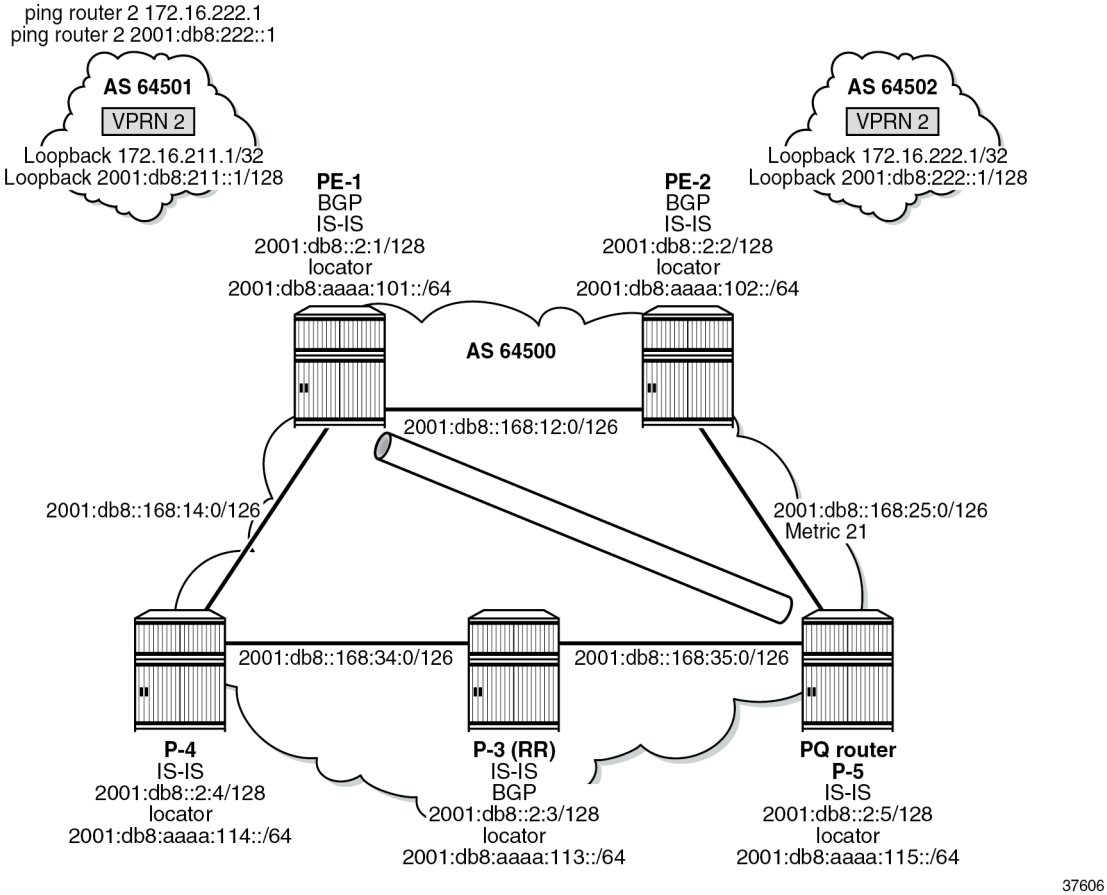
```
[/]
A:admin@PE-1# show router fp-tunnel-table 1 ipv6

===============================================================================
IPv6 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                 Protocol         Tunnel-ID
  Lbl/SID
    NextHop                                                  Intf/Tunnel
  Lbl/SID (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
2001:db8:aaaa:102::/64                      SRV6             524289
  -
    fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                     1/1/c1/1:1000
2001:db8:aaaa:113::/64                      SRV6             524290
  -
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"                      1/1/c2/1:1000
2001:db8:aaaa:114::/64                      SRV6             524291
```

```
  -
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"                          1/1/c2/1:1000
2001:db8:aaaa:115::/64                      SRV6                 524292
  -
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"                          1/1/c2/1:1000
-------------------------------------------------------------------------------
Total Entries : 4
-------------------------------------------------------------------------------
===============================================================================
```

Configure LFA on PE-1

Figure 66: Example topology with metric 21 between PE-2 and P-5 shows the example topology with initial metrics that is used to verify the behavior when a PQ-router provides TI-LFA protection.

Figure 66: Example topology with metric 21 between PE-2 and P-5



Configure regular LFA:

```
[/]
A:admin@PE-1# configure {
    router "Base" {
        isis 0 {
```

```
            loopfree-alternate {
            }
            exit all
```

Verify the IPv6 route table. There are two additional routes, corresponding with the End.X functions for locator “PE-1_loc” that are instantiated. The existing route to P-5 is loop-protected with regular LFA.

```
[/]
A:admin@PE-1# show router route-table ipv6

===============================================================================
IPv6 Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
---snip---
2001:db8::2:5/128 [L]                         Remote  ISIS      00h10m43s  18
       fe80::616:1ff:fe01:1-"int-PE-1-P-4"                          30
---snip---
2001:db8:aaaa:101:78a6:d000::/128             Local   ISIS      00h00m44s  18
       2001:db8:aaaa:101:78a6:d000:: (tunneled:SRV6-ISIS)           10
2001:db8:aaaa:101:78a6:e000::/128             Local   ISIS      00h00m44s  18
       2001:db8:aaaa:101:78a6:e000:: (tunneled:SRV6-ISIS)           10
---snip---
-------------------------------------------------------------------------------
No. of Routes: 18
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       ---snip---
===============================================================================
```

Verify the corresponding IPv6 FIB.

```
[/]
A:admin@PE-1# show router fib 1 ipv6

===============================================================================
FIB Display
===============================================================================
Prefix [Flags]                                          Protocol
  NextHop
-------------------------------------------------------------------------------
---snip---
2001:db8::2:5/128                                       ISIS
  fe80::616:1ff:fe01:1 (int-PE-1-P-4)
---snip---
2001:db8:aaaa:101:78a6:d000::/128                       ISIS
  2001:db8:aaaa:101:78a6:d000:: (Transport:SRV6:524293)
2001:db8:aaaa:101:78a6:e000::/128                       ISIS
  2001:db8:aaaa:101:78a6:e000:: (Transport:SRV6:524294)
---snip---
-------------------------------------------------------------------------------
Total Entries : 18
-------------------------------------------------------------------------------
===============================================================================
```

The IS-IS database contains additional information about the End.X functions that are instantiated on PE-1. The End.X functions for the locator “PE-1_loc” are instantiated and advertised. There are no changes for the other routers.

```
[/]
A:admin@PE-1# show router isis database detail

===============================================================================
Rtr Base ISIS Instance 0 Database (detail)
===============================================================================
---snip---
Displaying Level 2 database
-------------------------------------------------------------------------------
LSP ID    : PE-1.00-00                                  Level     : L2
---snip---
TLVs :
  ---snip---
  TE IS Nbrs   :
    Nbr   : PE-2.00
    Default Metric  : 10
    ---snip---
    End.X-SID: 2001:db8:aaaa:101:78a6:e000:: flags:B algo:0 weight:0 endpoint:End.X-USP
  TE IS Nbrs   :
    Nbr   : P-4.00
    Default Metric  : 10
    ---snip---
    End.X-SID: 2001:db8:aaaa:101:78a6:d000:: flags:B algo:0 weight:0 endpoint:End.X-USP
---snip---
Level (2) LSP Count : 5
---snip---
===============================================================================
```

Verify the SRv6 local SIDs and SRv6 base routing instance on PE-1. The End.X functions are also instantiated.

```
[/]
A:admin@PE-1# show router segment-routing-v6 local-sid

===============================================================================
Segment Routing v6 Local SIDs
===============================================================================
SID                                                 Type           Function
  Locator
  Context
-------------------------------------------------------------------------------
---snip---
2001:db8:aaaa:101:78a6:d000::                       End.X          494189
  PE-1_loc
    None
2001:db8:aaaa:101:78a6:e000::                       End.X          494190
  PE-1_loc
    None
-------------------------------------------------------------------------------
SIDs : 5
-------------------------------------------------------------------------------
===============================================================================
```

The SRv6 functions are listed.

```
[/]
A:admin@PE-1# show router segment-routing-v6 base-routing-instance
```

```
===============================================================================
Segment Routing v6 Base Routing Instance
===============================================================================
Locator
  Type       Function    SID                                      Status/InstId
    SRH-mode Protection  Interface
-------------------------------------------------------------------------------
PE-1_loc
  End                 1 2001:db8:aaaa:101:0:1000::              ok
    USP
-------------------------------------------------------------------------------
  Auto-allocated End.X: USP Protected,
-------------------------------------------------------------------------------
  End.X          *494189 2001:db8:aaaa:101:78a6:d000::          0
    USP        Protected int-PE-1-P-4
    ISIS Level: L2 Mac Address: 04:16:01:01:00:01 Nbr Sys Id: 0010.0100.1004
  End.X          *494190 2001:db8:aaaa:101:78a6:e000::          0
    USP        Protected int-PE-1-PE-2
    ISIS Level: L2 Mac Address: 04:0e:01:01:00:01 Nbr Sys Id: 0010.0100.1002
-------------------------------------------------------------------------------
===============================================================================
Legend: * - System allocated
```

Verify the IPv6 tunnel table. There are two new SRv6 tunnels for the End.X functions and the existing SRv6 tunnel to P-5 is loop-protected via regular LFA.

```
[/]
A:admin@PE-1# show router tunnel-table ipv6

===============================================================================
IPv6 Tunnel Table (Router: Base)
===============================================================================
Destination                                   Owner     Encap TunnelId  Pref
Nexthop                                       Color           Metric
-------------------------------------------------------------------------------
2001:db8:aaaa:101:78a6:d000::/128             srv6-isis SRV6  524293    0
  fe80::616:1ff:fe01:1-"int-PE-1-P-4"                           10
2001:db8:aaaa:101:78a6:e000::/128             srv6-isis SRV6  524294    0
  fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                          10
---snip---
2001:db8:aaaa:115::/64 [L]                    srv6-isis SRV6  524292    0
  fe80::616:1ff:fe01:1-"int-PE-1-P-4"                           30
-------------------------------------------------------------------------------
Flags: B = BGP or MPLS backup hop available
       L = Loop-Free Alternate (LFA) hop available
       ---snip---
===============================================================================
```

Verify the interfaces that the tunnels are using.

```
[/]
A:admin@PE-1# show router fp-tunnel-table 1 ipv6

===============================================================================
IPv6 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                 Protocol         Tunnel-ID
```

```
  Lbl/SID
    NextHop                                                 Intf/Tunnel
  Lbl/SID (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
---snip---
2001:db8:aaaa:115::/64                      SRV6              524292
  -
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"                     1/1/c2/1:1000
  -
    fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"(B)                 1/1/c1/1:1000
2001:db8:aaaa:101:78a6:d000::/128           SRV6              524293
  -
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"                     1/1/c2/1:1000
2001:db8:aaaa:101:78a6:e000::/128           SRV6              524294
  -
    fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                    1/1/c1/1:1000
-------------------------------------------------------------------------------
Total Entries : 6
-------------------------------------------------------------------------------
===============================================================================
```

Configure TI-LFA:

```
[/]
A:admin@PE-1# configure {
    router "Base" {
        isis 0 {
            loopfree-alternate {
                ti-lfa {
                }
            }
            exit all
```

There are no changes to the IPv6 route table, IPv6 FIB, IS-IS database, SRv6 local SIDs, and SRv6 base routing instance, while the change in LFA computation results in LFA protection for the tunnels to the remote routers. The existing SRv6 tunnels to PE-2, P-3 and P-4 are now also loop-protected. Verify the IPv6 tunnel table.

```
[/]
A:admin@PE-1# show router tunnel-table ipv6

===============================================================================
IPv6 Tunnel Table (Router: Base)
===============================================================================
Destination                                 Owner     Encap TunnelId  Pref
Nexthop                                     Color           Metric
-------------------------------------------------------------------------------
2001:db8:aaaa:101:78a6:d000::/128           srv6-isis SRV6  524293    0
  fe80::616:1ff:fe01:1-"int-PE-1-P-4"                         10
2001:db8:aaaa:101:78a6:e000::/128           srv6-isis SRV6  524294    0
  fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                        10
2001:db8:aaaa:102::/64 [L]                  srv6-isis SRV6  524289    0
  fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                        10
2001:db8:aaaa:113::/64 [L]                  srv6-isis SRV6  524290    0
  fe80::616:1ff:fe01:1-"int-PE-1-P-4"                         20
2001:db8:aaaa:114::/64 [L]                  srv6-isis SRV6  524291    0
  fe80::616:1ff:fe01:1-"int-PE-1-P-4"                         10
2001:db8:aaaa:115::/64 [L]                  srv6-isis SRV6  524292    0
  fe80::616:1ff:fe01:1-"int-PE-1-P-4"                         30
-------------------------------------------------------------------------------
Flags: B = BGP or MPLS backup hop available
```

```
       L = Loop-Free Alternate (LFA) hop available
       ---snip---
===============================================================================
```

Verify the interfaces that the tunnels are using. When the link between PE-1 and PE-2 is operational, SRv6 data is transported to PE-2 over this link, via next hop fe80::60e:1ff:fe01:1-"int-PE-1-PE-2". When the link between PE-1 and PE-2 fails, SRv6 data is transported to PE-2 using a fast reroute (FRR) backup link between PE-1 and P-4, via backup next hop fe80::616:1ff:fe01:1-"int-PE-1-P-4". The SRv6 data is transported to PE-2 then, via an SRv6 tunnel to the End function on P-5, as the backup SID 2001:db8:aaaa:115:0:1000:: indicates.

```
[/]
A:admin@PE-1# show router fp-tunnel-table 1 ipv6

===============================================================================
IPv6 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                Protocol         Tunnel-ID
  Lbl/SID
    NextHop                                               Intf/Tunnel
  Lbl/SID (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
2001:db8:aaaa:102::/64                     SRV6              524289
  -
    fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                    1/1/c1/1:1000
  2001:db8:aaaa:115:0:1000::
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"(B)                  1/1/c2/1:1000
2001:db8:aaaa:113::/64                     SRV6              524290
  -
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"                     1/1/c2/1:1000
  2001:db8:aaaa:115:0:1000::
    fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"(B)                 1/1/c1/1:1000
2001:db8:aaaa:114::/64                     SRV6              524291
  -
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"                     1/1/c2/1:1000
  2001:db8:aaaa:115:0:1000::
    fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"(B)                 1/1/c1/1:1000
2001:db8:aaaa:115::/64                     SRV6              524292
  -
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"                     1/1/c2/1:1000
  -
    fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"(B)                 1/1/c1/1:1000
2001:db8:aaaa:101:78a6:d000::/128          SRV6              524293
  -
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"                     1/1/c2/1:1000
2001:db8:aaaa:101:78a6:e000::/128          SRV6              524294
  -
    fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                    1/1/c1/1:1000
-------------------------------------------------------------------------------
Total Entries : 6
-------------------------------------------------------------------------------
===============================================================================
```

With the topology as shown in Figure 66: Example topology with metric 21 between PE-2 and P-5, this behavior is described as follows:

There is no regular LFA protection for the destination prefix to PE-2 using the protected PE-1-PE-2 link, which can be understood when the regular LFA inequality is determined using a shortest-path distance (Spd) calculation:

$$\text{Spd}(N, D) < \text{Spd}(N, S) + \text{Spd}(S, D)$$

where

Spd is the shortest path distance (according to level 2 metrics)

S is the source router (PE-1)

D is the destination router (PE-2)

N is the alternate next hop router or neighboring node (P-4)

If the outcome of the calculation is true, then regular LFA protection is valid; if the outcome is false, then there is no LFA protection.

In this case the outcome is false.

$$\text{Spd}(P\text{-}4, PE\text{-}2) < \text{Spd}(P\text{-}4, PE\text{-}1) + \text{Spd}(PE\text{-}1, PE\text{-}2)$$

$$(10 + 10 + 21) < 10 + 10$$

There is TI-LFA protection for all destination prefixes using the protected PE-1-PE-2 link, which is determined using the calculation for TI-LFA.

The TI-LFA inequality for the extended P-space P' is:

$$\text{Spd}(N, Y_i) < \text{Spd}(N, S) + \text{Spd}(S, Y_i)$$

$$\text{Spd}(P\text{-}4, Y_i) < \text{Spd}(P\text{-}4, PE\text{-}1) + \text{Spd}(PE\text{-}1, Y_i)$$

where Y_i is the set of routers {P-3, P-5} that are reachable from PE-1 and its neighbor P-4 on the post-convergence path to PE-2, without traversing the link between PE-1 and PE-2.

Apply this inequality to the set of routers Y_i:

For Y_i=P-3, the outcome is true. So P-3 is in P':

$$\text{Spd}(P\text{-}4, P\text{-}3) < \text{Spd}(P\text{-}4, PE\text{-}1) + \text{Spd}(PE\text{-}1, P\text{-}3)$$

$$10 < 10 + (10 + 10)$$

For Y_i=P-5, the outcome is true. So P-5 is in P':

$$\text{Spd}(P\text{-}4, P\text{-}5) < \text{Spd}(P\text{-}4, PE\text{-}1) + \text{Spd}(PE\text{-}1, P\text{-}5)$$

$$(10 + 10) < 10 + (10 + 10 + 10)$$

So, the extended P-space P' = {P-3, P-5}

The TI-LFA inequality for the Q-space Q is:

$$\text{Spd}(Z_i, D) < \text{Spd}(Z_i, S) + \text{Spd}(S, D)$$

$$Spd(Z_i, PE\text{-}2) < Spd(Z_i, PE\text{-}1) + Spd(PE\text{-}1, PE\text{-}2)$$

where Z_i is the set of routers {P-3, P-5} that are reachable from PE-2 using reverse SPF on the post-convergence path to PE-1 without traversing the link between PE-1 and PE-2.

Apply this inequality to the set of routers Z_i:

For Z_i=P-3, the outcome is false. So P-3 is **not** in Q:

$$Spd(P\text{-}3, PE\text{-}2) < Spd(P\text{-}3, PE\text{-}1) + Spd(PE\text{-}1, PE\text{-}2)$$

$$(10 + 21) < (10 + 10) + 10$$

For Z_i=P-5, the outcome is true. So P-5 is in Q:

$$Spd(P\text{-}5, PE\text{-}2) < Spd(P\text{-}5, PE\text{-}1) + Spd(PE\text{-}1, PE\text{-}2)$$

$$21 < (10 + 10 + 10) + 10$$

So, the Q-space Q = {P-5}

So, the link between PE-1 and PE-2 is TI-LFA protected with the PQ-router P-5 that belongs to the intersection of the extended P space P’ and the Q space.

SRv6 data is transported to P-4, P-3, and P-5 over the link between PE-1 and P-4, via next hop fe80::616:1ff:fe01:1-"int-PE-1-P-4". When the link between PE-1 and P-4 fails, SRv6 data is transported to P-4, P-3, and P-5 using a FRR backup link between PE-1 and PE-2, via backup next hop fe80::60e:1ff:fe01:1-"int-PE-1-PE-2". The SRv6 data is transported to P-4 and P-3 then via an SRv6 tunnel to the End function on P-5, as the backup SID 2001:db8:aaaa:115:0:1000:: indicates. The SRv6 data is transported to P-5 then without using an SRv6 tunnel, as the absence of a backup SID indicates.

Disable the link between PE-1 and PE-2:

```
[/]
A:admin@PE-1# configure {
    router "Base" {
        interface "int-PE-1-PE-2" {
            admin-state disable
        }
        exit all
```

Because PE-2 disappears as a traffic-engineered (TE) IS-IS neighbor of PE-1, the End.X function that corresponds with the interface “int-PE-1-PE-2” is no longer instantiated. The IPv6 route table and IPv6 FIB indicate that data transport from PE-1 to PE-2 and P-5 now follows a path with a higher metric via P-4. The route to the End.X function that corresponds with the interface “int-PE-1-PE-2” is no longer present. There is no longer LFA protection for the route to P-5. Verify the IPv6 route table.

```
[/]
A:admin@PE-1# show router route-table ipv6

===============================================================================
IPv6 Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
---snip---
```

```
2001:db8::2:2/128                                Remote  ISIS      00h00m46s  18
       fe80::616:1ff:fe01:1-"int-PE-1-P-4"                            51
---snip---
2001:db8::2:5/128                                Remote  ISIS      00h16m33s  18
       fe80::616:1ff:fe01:1-"int-PE-1-P-4"                            30
2001:db8::168:12:0/126                           Remote  ISIS      00h00m46s  18
       fe80::616:1ff:fe01:1-"int-PE-1-P-4"                            61
---snip---
2001:db8::168:25:0/126                           Remote  ISIS      00h00m46s  18
       fe80::616:1ff:fe01:1-"int-PE-1-P-4"                            51
---snip---
-------------------------------------------------------------------------------
No. of Routes: 17
---snip---
===============================================================================
```

Verify the corresponding IPv6 FIB.

```
[/]
A:admin@PE-1# show router fib 1 ipv6

===============================================================================
FIB Display
===============================================================================
Prefix [Flags]                                                 Protocol
  NextHop
-------------------------------------------------------------------------------
---snip---
2001:db8::2:2/128                                              ISIS
  fe80::616:1ff:fe01:1 (int-PE-1-P-4)
---snip---
2001:db8::168:12:0/126                                         ISIS
  fe80::616:1ff:fe01:1 (int-PE-1-P-4)
---snip---
2001:db8::168:25:0/126                                         ISIS
  fe80::616:1ff:fe01:1 (int-PE-1-P-4)
---snip---
-------------------------------------------------------------------------------
Total Entries : 17
-------------------------------------------------------------------------------
===============================================================================
```

Verify the SRv6 local SIDs and SRv6 base routing instance on PE-1. The SID that corresponds with the interface “int-PE-1-PE-2” is no longer present and is no longer advertised to the other routers.

```
[/]
A:admin@PE-1# show router segment-routing-v6 local-sid

===============================================================================
Segment Routing v6 Local SIDs
===============================================================================
SID                                             Type           Function
  Locator
  Context
-------------------------------------------------------------------------------
2001:db8:aaaa:101:0:1000::                      End            1
  PE-1_loc
  Base
2001:db8:aaaa:101:78a6:b000::                   End.DT6        494187
  PE-1_loc
  SvcId: 2 Name: VPRN_2
2001:db8:aaaa:101:78a6:c000::                   End.DT4        494188
```

```
  PE-1_loc
  SvcId: 2 Name: VPRN_2
2001:db8:aaaa:101:78a6:d000::                    End.X          494189
  PE-1_loc
    None
-------------------------------------------------------------------------------
SIDs : 4
-------------------------------------------------------------------------------
===============================================================================
```

The End.X function with SID 2001:db8:aaaa:101:78a6:e000:: that corresponds with the interface “int-PE-1-PE-2” is no longer instantiated.

```
[/]
A:admin@PE-1# show router segment-routing-v6 base-routing-instance

===============================================================================
Segment Routing v6 Base Routing Instance
===============================================================================
Locator
  Type       Function    SID                                      Status/InstId
    SRH-mode Protection  Interface
-------------------------------------------------------------------------------
PE-1_loc
  End                 1 2001:db8:aaaa:101:0:1000::              ok
    USP
-------------------------------------------------------------------------------
  Auto-allocated End.X: USP Protected,
-------------------------------------------------------------------------------
  End.X          *494189 2001:db8:aaaa:101:78a6:d000::          0
    USP        Protected int-PE-1-P-4
    ISIS Level: L2 Mac Address: 04:16:01:01:00:01 Nbr Sys Id: 0010.0100.1004
-------------------------------------------------------------------------------
===============================================================================
Legend: * - System allocated
```

Verify the IPv6 tunnel table. There are no longer any backup tunnels and SRv6 data is transported to all destinations via the link between PE-1 and P-4.

```
[/]
A:admin@PE-1# show router tunnel-table ipv6

===============================================================================
IPv6 Tunnel Table (Router: Base)
===============================================================================
Destination                                 Owner     Encap TunnelId  Pref
Nexthop                                     Color           Metric
-------------------------------------------------------------------------------
2001:db8:aaaa:101:78a6:d000::/128           srv6-isis SRV6  524293    0
  fe80::616:1ff:fe01:1-"int-PE-1-P-4"                         10
2001:db8:aaaa:102::/64                      srv6-isis SRV6  524289    0
  fe80::616:1ff:fe01:1-"int-PE-1-P-4"                         51
2001:db8:aaaa:113::/64                      srv6-isis SRV6  524290    0
  fe80::616:1ff:fe01:1-"int-PE-1-P-4"                         20
2001:db8:aaaa:114::/64                      srv6-isis SRV6  524291    0
  fe80::616:1ff:fe01:1-"int-PE-1-P-4"                         10
2001:db8:aaaa:115::/64                      srv6-isis SRV6  524292    0
  fe80::616:1ff:fe01:1-"int-PE-1-P-4"                         30
-------------------------------------------------------------------------------
---snip---
===============================================================================
```

Verify the interfaces that the tunnels are using. There is no longer any possibility for alternate routes.

```
[/]
A:admin@PE-1# show router fp-tunnel-table 1 ipv6

================================================================================
IPv6 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
================================================================================
Destination                                Protocol         Tunnel-ID
  Lbl/SID
    NextHop                                               Intf/Tunnel
  Lbl/SID (backup)
    NextHop   (backup)
--------------------------------------------------------------------------------
2001:db8:aaaa:102::/64                     SRV6              524289
  -
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"                    1/1/c2/1:1000
2001:db8:aaaa:113::/64                     SRV6              524290
  -
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"                    1/1/c2/1:1000
2001:db8:aaaa:114::/64                     SRV6              524291
  -
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"                    1/1/c2/1:1000
2001:db8:aaaa:115::/64                     SRV6              524292
  -
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"                    1/1/c2/1:1000
2001:db8:aaaa:101:78a6:d000::/128          SRV6              524293
  -
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"                    1/1/c2/1:1000
--------------------------------------------------------------------------------
Total Entries : 5
--------------------------------------------------------------------------------
================================================================================
```

Enable the link between PE-1 and PE-2 to restore the initial topology:

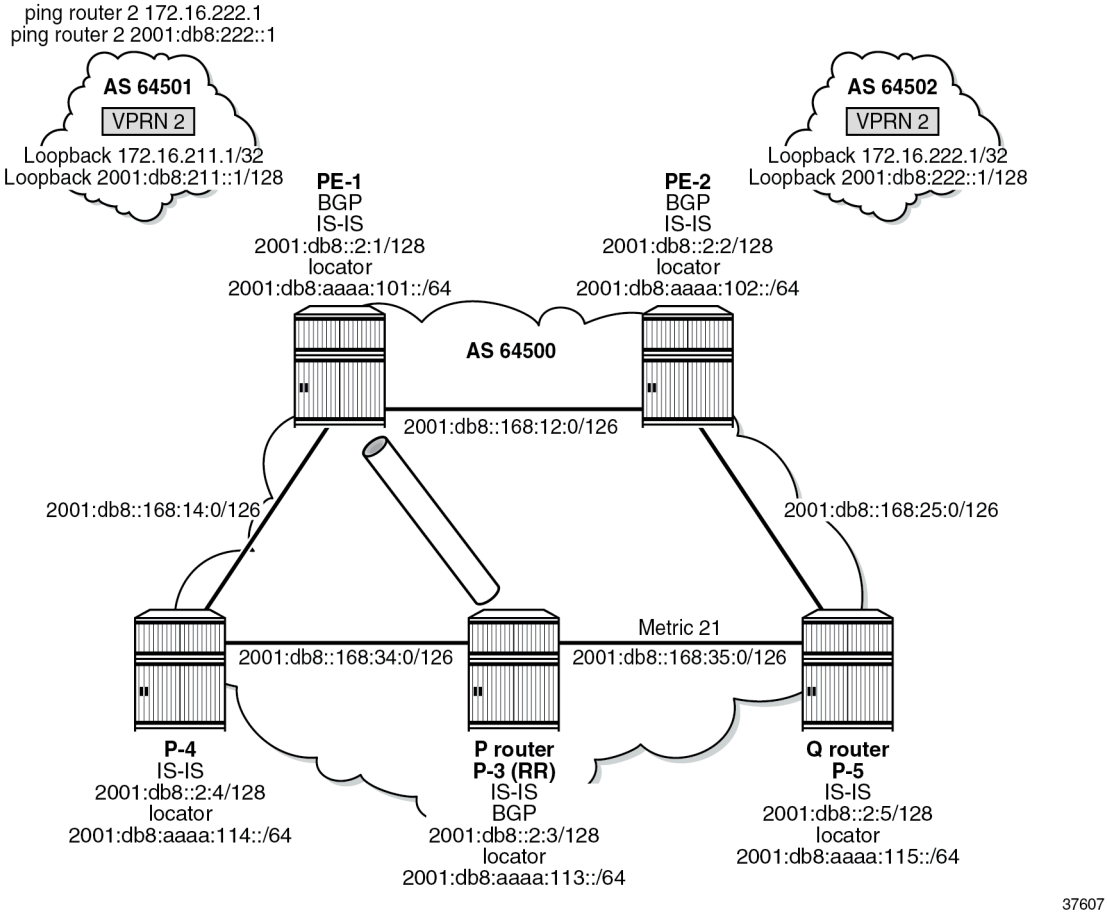
```
[/]
A:admin@PE-1# configure {
    router "Base" {
        interface "int-PE-1-PE-2" {
            admin-state enable
        }
        exit all
```

The End.X function that corresponds with the interface “int-PE-1-PE-2” is re-instantiated, but with SID 2001:db8:aaaa:101:78a6:f000:: and SRv6 Tunnel-ID 524295.

Modify metrics so that the P-router and the Q-router no longer coincide

Figure 3 shows the example topology with modified metrics that is used to verify the behavior when a disjointed P-router and Q-router provide TI-LFA protection.

Figure 67: Example topology with metric 21 between P-3 and P-5



Metrics can be modified for the interface "int-PE-2-P-5" on PE-2 with the command **configure router "Base" isis 0 interface "int-PE-2-P-5" level 2 metric <value>**. Similar commands apply for the interface "int-P-3-P-5" on P-3, and for the interfaces "int-P-5-PE-2" and "int-P-5-P-3" on P-5.

Verify the IPv6 route table. P-5 is now reached via interface “int-PE-1-PE-2”.

```
[/]
A:admin@PE-1# show router route-table ipv6

===============================================================================
IPv6 Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
---snip---
2001:db8::2:5/128                             Remote  ISIS      00h01m07s  18
       fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                         20
---snip---
2001:db8::168:35:0/126                        Remote  ISIS      00h00m31s  18
       fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                         41
---snip---
-------------------------------------------------------------------------------
```

```
No. of Routes: 18
---snip---
===============================================================================
```

Verify the corresponding IPv6 FIB.

```
[/]
A:admin@PE-1# show router fib 1 ipv6

===============================================================================
FIB Display
===============================================================================
Prefix [Flags]                                              Protocol
  NextHop
-------------------------------------------------------------------------------
---snip---
2001:db8::2:5/128                                           ISIS
  fe80::60e:1ff:fe01:1 (int-PE-1-PE-2)
---snip---
2001:db8::168:35:0/126                                      ISIS
  fe80::60e:1ff:fe01:1 (int-PE-1-PE-2)
---snip---
-------------------------------------------------------------------------------
Total Entries : 18
-------------------------------------------------------------------------------
===============================================================================
```

On PE-1, apart from the metrics, the IS-IS data base, the SRv6 local SIDs, and the SRv6 base routing instance do not change.

```
[/]
A:admin@PE-1# show router tunnel-table ipv6

===============================================================================
IPv6 Tunnel Table (Router: Base)
===============================================================================
Destination                                 Owner     Encap TunnelId  Pref
Nexthop                                     Color           Metric
-------------------------------------------------------------------------------
---snip---
2001:db8:aaaa:115::/64                      srv6-isis SRV6  524292    0
  fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                        20
-------------------------------------------------------------------------------
---snip---
===============================================================================
```

Verify the interfaces that the tunnels are using.

```
[/]
A:admin@PE-1# show router fp-tunnel-table 1 ipv6

===============================================================================
IPv6 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                 Protocol         Tunnel-ID
  Lbl/SID
    NextHop                                                  Intf/Tunnel
  Lbl/SID (backup)
```

```
    NextHop   (backup)
-------------------------------------------------------------------------------
---snip---
2001:db8:aaaa:115::/64                      SRV6             524292
  -
    fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                     1/1/c1/1:1000
---snip---
-------------------------------------------------------------------------------
Total Entries : 6
-------------------------------------------------------------------------------
===============================================================================
```

Without support for LFA on P-3, the TI-LFA computation on PE-1 does not lead to a PQ-router, because the End.X functions on P-3 are neither instantiated nor advertised to the other routers.

On P-3, verify the SRv6 local SIDs and SRv6 base routing instance. The End.X functions are not yet instantiated.

```
[/]
A:admin@P-3# show router segment-routing-v6 local-sid

===============================================================================
Segment Routing v6 Local SIDs
===============================================================================
SID                                             Type           Function
  Locator
  Context
-------------------------------------------------------------------------------
2001:db8:aaaa:113:0:1000::                      End            1
  P-3_loc
  Base
-------------------------------------------------------------------------------
SIDs : 1
-------------------------------------------------------------------------------
===============================================================================
```

Only the End function is already instantiated. The End.X functions are not yet instantiated.

```
[/]
A:admin@P-3# show router segment-routing-v6 base-routing-instance

===============================================================================
Segment Routing v6 Base Routing Instance
===============================================================================
Locator
  Type       Function    SID                                     Status/InstId
    SRH-mode Protection  Interface
-------------------------------------------------------------------------------
P-3_loc
  End                 1 2001:db8:aaaa:113:0:1000::              ok
    USP
-------------------------------------------------------------------------------
  Auto-allocated End.X: USP Protected,
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
===============================================================================
Legend: * - System allocated
```

Configure LFA on P-3

```
[/]
A:admin@P-3# configure {
    router "Base" {
        isis 0 {
            loopfree-alternate {
            }
            exit all
```

The IS-IS database contains additional information about the End.X functions that are instantiated on P-3.

```
[/]
A:admin@PE-1# show router isis database detail

===============================================================================
Rtr Base ISIS Instance 0 Database (detail)
===============================================================================
---snip---
Displaying Level 2 database
-------------------------------------------------------------------------------
---snip---
-------------------------------------------------------------------------------
LSP ID    : P-3.00-00                                   Level     : L2
---snip---
TLVs :
  ---snip---
  TE IS Nbrs   :
    Nbr   : P-4.00
    Default Metric  : 10
    ---snip---
    End.X-SID: 2001:db8:aaaa:113:0:4000:: flags:B algo:0 weight:0 endpoint:End.X-USP
  TE IS Nbrs   :
    Nbr   : P-5.00
    Default Metric  : 21
    ---snip---
    End.X-SID: 2001:db8:aaaa:113:0:5000:: flags:B algo:0 weight:0 endpoint:End.X-USP
---snip---
Level (2) LSP Count : 5
-------------------------------------------------------------------------------
---snip--
===============================================================================
```

On PE-1, the IS-IS data base, the IPv6 route table, the IPv6 FIB, the SRv6 local SIDs, and the SRv6 base routing instance do not change.

On PE-3, verify the SRv6 local SIDs and SRv6 base routing instance. The End.X functions are also instantiated.

```
[/]
A:admin@P-3# show router segment-routing-v6 local-sid

===============================================================================
Segment Routing v6 Local SIDs
===============================================================================
SID                                                 Type           Function
  Locator
  Context
-------------------------------------------------------------------------------
2001:db8:aaaa:113:0:1000::                          End            1
```

```
  P-3_loc
  Base
2001:db8:aaaa:113:0:4000::                          End.X          4
  P-3_loc
    None
2001:db8:aaaa:113:0:5000::                          End.X          5
  P-3_loc
    None
-------------------------------------------------------------------------------
SIDs : 3
-------------------------------------------------------------------------------
===============================================================================
```

```
[/]
A:admin@P-3# show router segment-routing-v6 base-routing-instance

===============================================================================
Segment Routing v6 Base Routing Instance
===============================================================================
Locator
  Type       Function    SID                                      Status/InstId
    SRH-mode Protection  Interface
-------------------------------------------------------------------------------
P-3_loc
  End                  1 2001:db8:aaaa:113:0:1000::               ok
    USP
-------------------------------------------------------------------------------
  Auto-allocated End.X: USP Protected,
-------------------------------------------------------------------------------
  End.X               *4 2001:db8:aaaa:113:0:4000::               0
    USP        Protected int-P-3-P-4
    ISIS Level: L2 Mac Address: 04:16:01:01:00:0b Nbr Sys Id: 0010.0100.1004
  End.X               *5 2001:db8:aaaa:113:0:5000::               0
    USP        Protected int-P-3-P-5
    ISIS Level: L2 Mac Address: 04:1a:01:01:00:0b Nbr Sys Id: 0010.0100.1005
-------------------------------------------------------------------------------
===============================================================================
Legend: * - System allocated
```

Verify the IPv6 tunnel table. The existing routes to PE-2 and P-5 are loop-protected.

```
[/]
A:admin@PE-1# show router tunnel-table ipv6

===============================================================================
IPv6 Tunnel Table (Router: Base)
===============================================================================
Destination                                 Owner     Encap TunnelId  Pref
Nexthop                                     Color           Metric
-------------------------------------------------------------------------------
---snip---
2001:db8:aaaa:102::/64 [L]                  srv6-isis SRV6  524289    0
  fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                        10
---snip---
2001:db8:aaaa:115::/64 [L]                  srv6-isis SRV6  524292    0
  fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                        20
-------------------------------------------------------------------------------
Flags: B = BGP or MPLS backup hop available
       L = Loop-Free Alternate (LFA) hop available
       ---snip---
===============================================================================
```

Verify the interfaces that the tunnels are using. Interface “int-PE-1-PE-2” is configured on port 1/1/c1/1:1000. Interface “int-PE-1-P-4” is configured on port 1/1/c2/1:1000.

When the link between PE-1 and PE-2 is operational, SRv6 data is transported to PE-2 over this link, via next hop fe80::60e:1ff:fe01:1-"int-PE-1-PE-2". When the link between PE-1 and PE-2 fails, SRv6 data is transported to PE-2 using a FRR backup link between PE-1 and P-4, via backup next hop fe80::616:1ff:fe01:1-"int-PE-1-P-4". The SRv6 data is transported to PE-2 then via an SRv6 tunnel to the End.X function on P-3, as the backup SID 2001:db8:aaaa:113:0:5000:: indicates, followed by source routing to P-5.

```
[/]
A:admin@PE-1# show router fp-tunnel-table 1 ipv6

================================================================================
IPv6 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
================================================================================
Destination                                Protocol         Tunnel-ID
  Lbl/SID
    NextHop                                               Intf/Tunnel
  Lbl/SID (backup)
    NextHop   (backup)
--------------------------------------------------------------------------------
2001:db8:aaaa:102::/64                     SRV6              524289
  -
    fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                   1/1/c1/1:1000
  2001:db8:aaaa:113:0:5000::
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"(B)                 1/1/c2/1:1000
---snip---
2001:db8:aaaa:115::/64                     SRV6              524292
  -
    fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"                   1/1/c1/1:1000
  2001:db8:aaaa:113:0:5000::
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"(B)                 1/1/c2/1:1000
---snip---
--------------------------------------------------------------------------------
Total Entries : 6
--------------------------------------------------------------------------------
================================================================================
```

With the topology as shown in Figure 67: Example topology with metric 21 between P-3 and P-5, this behavior is described as follows:

There is no regular LFA protection for the destination prefix to PE-2 using the protected PE-1-PE-2 link, which can be understood when the regular LFA inequality is determined using a shortest-path distance (Spd) calculation:

$$\text{Spd}(N, D) < \text{Spd}(N, S) + \text{Spd}(S, D)$$

where

Spd is the shortest path distance (according to level 2 metrics)

S is the source router (PE-1)

D is the destination router (PE-2)

N is the alternate next hop router or neighboring node (P-4)

If the outcome of the calculation is true, then regular LFA protection is valid; if the outcome is false, then there is no LFA protection.

In this case the outcome is false.

Spd(*P-4*, *PE-2*) < Spd(*P-4*, *PE-1*) + Spd(*PE-1*, *PE-2*)

(10 + 21 + 10) < 10 + 10

There is TI-LFA protection for all destination prefixes using the protected PE-1-PE-2 link:

The TI-LFA inequality for the extended P-space P' is:

Spd(N, Y_i) < Spd(N, S) + Spd(S, Y_i)

Spd(*P-4*, Y_i) < Spd(*P-4*, *PE-1*) + Spd(*PE-1*, Y_i)

where Y_i is the set of routers {P-3, P-5} that are reachable from PE-1 and its neighbor P-4 on the post-convergence path to PE-2, without traversing the link between PE-1 and PE-2.

Apply this inequality to the set of routers Y_i:

For Y_i=P-3, the outcome is true. So P-3 is in P':

Spd(*P-4*, *P-3*) < Spd(*P-4*, *PE-1*) + Spd(*PE-1*, *P-3*)

10 < 10 + (10 + 10)

For Y_i=P-5, the outcome is false. So P-5 is **not** in P':

Spd(*P-4*, *P-5*) < Spd(*P-4*, *PE-1*) + Spd(*PE-1*, *P-5*)

(10 + 21) < 10 + (10 + 10)

So, the extended P-space P' = {P-3}

The TI-LFA inequality for the Q-space Q is:

Spd(Z_i, D) < Spd(Z_i, S) + Spd(S, D)

Spd(Z_i, *PE-2*) < Spd(Z_i, *PE-1*) + Spd(*PE-1*, *PE-2*)

where Z_i is the set of routers {P-3, P-5} that are reachable from PE-2 using reverse SPF on the post-convergence path to PE-1 without traversing the link between PE-1 and PE-2.

Apply this inequality to the set of routers Z_i:

For Z_i=P-3, the outcome is false. So P-3 is **not** in Q:

Spd(*P-3*, *PE-2*) < Spd(*P-3*, *PE-1*) + Spd(*PE-1*, *PE-2*)

(21 + 10) < (10 + 10) + 10

For Z_i=P-5, the outcome is true. So P-5 is in Q:

Spd(*P-5*, *PE-2*) < Spd(*P-5*, *PE-1*) + Spd(*PE-1*, *PE-2*)

10 < (21 + 10 + 10) + 10

So, the Q-space Q = {P-5}

So, the link between PE-1 and PE-2 is TI-LFA protected with the P-router P-3 and Q-router P-5.

Disable the link between PE-1 and PE-2:

```
[/]
A:admin@PE-1# configure {
    router "Base" {
        interface "int-PE-1-PE-2" {
            admin-state disable
        }
        exit all
```

Because PE-2 disappears as a TE IS-IS neighbor of PE-1, the End.X function that corresponds with the interface “int-PE-1-PE-2” is no longer instantiated. The IPv6 route table and IPv6 FIB indicate that data transport from PE-1 to PE-2 and P-5 now follows a path with a higher metric via P-4. Verify the IPv6 route table.

```
[/]
A:admin@PE-1# show router route-table ipv6

===============================================================================
IPv6 Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
---snip---
2001:db8::2:2/128                             Remote  ISIS      00h00m23s  18
       fe80::616:1ff:fe01:1-"int-PE-1-P-4"                          51
---snip---
2001:db8::2:5/128                             Remote  ISIS      00h00m23s  18
       fe80::616:1ff:fe01:1-"int-PE-1-P-4"                          41
2001:db8::168:12:0/126                        Remote  ISIS      00h00m23s  18
       fe80::616:1ff:fe01:1-"int-PE-1-P-4"                          61
---snip---
2001:db8::168:25:0/126                        Remote  ISIS      00h00m23s  18
       fe80::616:1ff:fe01:1-"int-PE-1-P-4"                          51
---snip---
2001:db8::168:35:0/126                        Remote  ISIS      00h00m23s  18
       fe80::616:1ff:fe01:1-"int-PE-1-P-4"                          41
---snip---
-------------------------------------------------------------------------------
No. of Routes: 17
---snip---
===============================================================================
```

Verify the corresponding IPv6 FIB.

```
[/]
A:admin@PE-1# show router fib 1 ipv6

===============================================================================
FIB Display
===============================================================================
Prefix [Flags]                                          Protocol
```

```
  NextHop
-------------------------------------------------------------------------------
---snip---
2001:db8::2:2/128                                          ISIS
  fe80::616:1ff:fe01:1 (int-PE-1-P-4)
---snip---
2001:db8::2:5/128                                          ISIS
  fe80::616:1ff:fe01:1 (int-PE-1-P-4)
2001:db8::168:12:0/126                                     ISIS
  fe80::616:1ff:fe01:1 (int-PE-1-P-4)
---snip---
2001:db8::168:25:0/126                                     ISIS
  fe80::616:1ff:fe01:1 (int-PE-1-P-4)
---snip---
2001:db8::168:35:0/126                                     ISIS
  fe80::616:1ff:fe01:1 (int-PE-1-P-4)
---snip---
-------------------------------------------------------------------------------
Total Entries : 17
-------------------------------------------------------------------------------
===============================================================================
```

Verify the SRv6 local SIDs and SRv6 base routing instance on PE-1. The SID that corresponds with the interface “int-PE-1-PE-2” is no longer present and is no longer advertised to the other routers.

```
[/]
A:admin@PE-1# show router segment-routing-v6 local-sid

===============================================================================
Segment Routing v6 Local SIDs
===============================================================================
SID                                             Type           Function
  Locator
  Context
-------------------------------------------------------------------------------
2001:db8:aaaa:101:0:1000::                      End            1
  PE-1_loc
  Base
2001:db8:aaaa:101:78a6:b000::                   End.DT6        494187
  PE-1_loc
  SvcId: 2 Name: VPRN_2
2001:db8:aaaa:101:78a6:c000::                   End.DT4        494188
  PE-1_loc
  SvcId: 2 Name: VPRN_2
2001:db8:aaaa:101:78a6:d000::                   End.X          494189
  PE-1_loc
    None
-------------------------------------------------------------------------------
SIDs : 4
-------------------------------------------------------------------------------
===============================================================================
```

The End.X function that corresponds with the interface “int-PE-1-PE-2” is no longer instantiated.

```
[/]
A:admin@PE-1# show router segment-routing-v6 base-routing-instance

===============================================================================
Segment Routing v6 Base Routing Instance
===============================================================================
Locator
  Type       Function    SID                                     Status/InstId
```

```
    SRH-mode Protection  Interface
-------------------------------------------------------------------------------
PE-1_loc
  End                 1 2001:db8:aaaa:101:0:1000::             ok
    USP
-------------------------------------------------------------------------------
  Auto-allocated End.X: USP Protected,
-------------------------------------------------------------------------------
  End.X          *494189 2001:db8:aaaa:101:78a6:d000::          0
    USP        Protected int-PE-1-P-4
    ISIS Level: L2 Mac Address: 04:16:01:01:00:01 Nbr Sys Id: 0010.0100.1004
-------------------------------------------------------------------------------
===============================================================================
Legend: * - System allocated
```

Verify the IPv6 tunnel table. There are no longer any backup tunnels and SRv6 data is transported to all destinations via the link between PE-1 and P-4.

```
[/]
A:admin@PE-1# show router tunnel-table ipv6

===============================================================================
IPv6 Tunnel Table (Router: Base)
===============================================================================
Destination                                     Owner     Encap TunnelId  Pref
Nexthop                                         Color           Metric
-------------------------------------------------------------------------------
2001:db8:aaaa:101:78a6:d000::/128               srv6-isis SRV6  524293    0
  fe80::616:1ff:fe01:1-"int-PE-1-P-4"                         10
2001:db8:aaaa:102::/64                          srv6-isis SRV6  524289    0
  fe80::616:1ff:fe01:1-"int-PE-1-P-4"                         51
2001:db8:aaaa:113::/64                          srv6-isis SRV6  524290    0
  fe80::616:1ff:fe01:1-"int-PE-1-P-4"                         20
2001:db8:aaaa:114::/64                          srv6-isis SRV6  524291    0
  fe80::616:1ff:fe01:1-"int-PE-1-P-4"                         10
2001:db8:aaaa:115::/64                          srv6-isis SRV6  524292    0
  fe80::616:1ff:fe01:1-"int-PE-1-P-4"                         41
-------------------------------------------------------------------------------
---snip---
===============================================================================
```

Verify the interfaces that the tunnels are using. There is no longer any possibility for alternate routes.

```
[/]
A:admin@PE-1# show router fp-tunnel-table 1 ipv6

===============================================================================
IPv6 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                 Protocol         Tunnel-ID
  Lbl/SID
    NextHop                                                  Intf/Tunnel
  Lbl/SID (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
2001:db8:aaaa:102::/64                      SRV6              524289
  -
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"                      1/1/c2/1:1000
```

```
2001:db8:aaaa:113::/64                        SRV6             524290
  -
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"                        1/1/c2/1:1000
2001:db8:aaaa:114::/64                        SRV6             524291
  -
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"                        1/1/c2/1:1000
2001:db8:aaaa:115::/64                        SRV6             524292
  -
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"                        1/1/c2/1:1000
2001:db8:aaaa:101:78a6:d000::/128             SRV6             524293
  -
    fe80::616:1ff:fe01:1-"int-PE-1-P-4"                        1/1/c2/1:1000
-------------------------------------------------------------------------------
Total Entries : 5
-------------------------------------------------------------------------------
===============================================================================
```

Enable the link between PE-1 and PE-2 to restore the initial topology:

```
[/]
A:admin@PE-1# configure {
    router "Base" {
        interface "int-PE-1-PE-2" {
            admin-state enable
        }
        exit all
```

The End.X function that corresponds with the interface “int-PE-1-PE-2” is re-instantiated, but with SID 2001:db8:aaaa:101:78a6:e000:: and Tunnel-ID 524296.

Conclusion

To guard against the failure of the initial data path, LFA protection via an LFA backup path is possible for SRv6 data transport.

SRv6 Policy Support for Layer 2 and Layer 3 Services

This chapter provides information about SRv6 policy support for Layer 2 and Layer 3 services.

Topics in this chapter include:

Applicability

SRv6 policies for Layer 2 and Layer 3 services are supported in SR OS Release 22.7.R1 and later. The information and configuration in this chapter are based on SR OS Release 24.3.R2.

MPLS Segment Routing (SR) policies are described in the BGP Signaled Segment Routing Policy chapter.

Overview

SRv6 policies are Segment Routing (SR) policies with an IPv6 data plane. SRv6 policies consist of one or more lists of SRv6 segments, where each segment list represents a source route that can be used to enable traffic engineering through SRv6 networks. Each segment list contains a list of SRv6 SIDs comprising the top Segment Identifier (SID) and the further SIDs. Both 128-bit classic SIDs and 16-bit micro-segment SIDs can be used and it is possible to combine SRv6 segments derived from classic SRv6 and micro-segment SRv6 locators. The combination of classic and micro-segment SIDs can happen at the policy level and at the segment list level. A policy can have a segment list with only classic SIDs and another segment list with only micro-segment SIDs. A segment list can itself contain both classic and micro-segment SIDs. Different types of SIDs can be used, such as node SIDs (for classic SIDs: End SIDs; for micro-segment SIDs: uN SIDs) or adjacency SIDs (End.X SIDs; uA SIDs).

An SRv6 policy is identified through the { head-end, endpoint, color } tuple:

- The head-end is the node where the SRv6 policy is instantiated. The head-end steers the traffic into the policy with the SID stack. From the perspective of the head-end, the SRv6 policy can be identified using the { color, endpoint } tuple.
- The endpoint is the IPv6 destination of the SRv6 policy.
- The color is a numerical value that is used by the head-end to associate the SRv6 policy with a characteristic, such as low-latency or high-throughput. The color is a 32-bit transitive extended community that forms part of the BGP NLRI which is exported by the endpoint, and the head-end resolves the next hop to that endpoint via an SR policy with the corresponding color.

Static SRv6 policies are configured through CLI, using the **configure router segment-routing sr-policies static-policy <..>** command with the **type srv6** option. Static SRv6 policies can be configured on the

head-end (**head-end local**) and it is also possible to configure static SRv6 policies on a remote node and signal these SRv6 policies using BGP to the head-end (**head-end <non-local IPv4 system address>**).

SRv6 policies are programmed in the IPv6 Tunnel Table Manager (TTMv6) and can be bound to BGP-based SRv6 services, such as IP VPN and EVPN services:

- EVPN VPWS
- EVPN VPLS
- EVPN IFL
- VPN IPv4
- VPN IPv6

Similar to SR policies with an MPLS data plane, the next-hop resolution is based on the color and on the comparison of the next hop with the SR policy endpoint. The head-end, endpoint, and color define a matching policy, but a matching policy can be an SR MPLS policy or an SRv6 policy. Therefore, the service also uses the data plane technology of the policy to select the tunnel type to resolve over. SRv6 services can only resolve over SRv6 tunnels and MPLS services cannot resolve over SRv6 tunnels. An SRv6 service cannot fall back to an MPLS tunnel type.

The **resolution** command controls the automatic binding of SRv6 services to SRv6 tunnels. The resolution options are as follows:

- **tunnel-table**

  The tunnel-table command option resolves the route directly to a tunnel in TTMv6. The system tries to find an SRv6 policy with the same endpoint and color for BGP routes received with an SRv6 TLV and that contain an SRv6 service SID in the TTMv6. If no such SRv6 policy is found in the TTMv6, the resolution fails.

- **route-table**

  The route-table command option is the default behavior which resolves the route to a shortest path SRv6 tunnel in the route table.

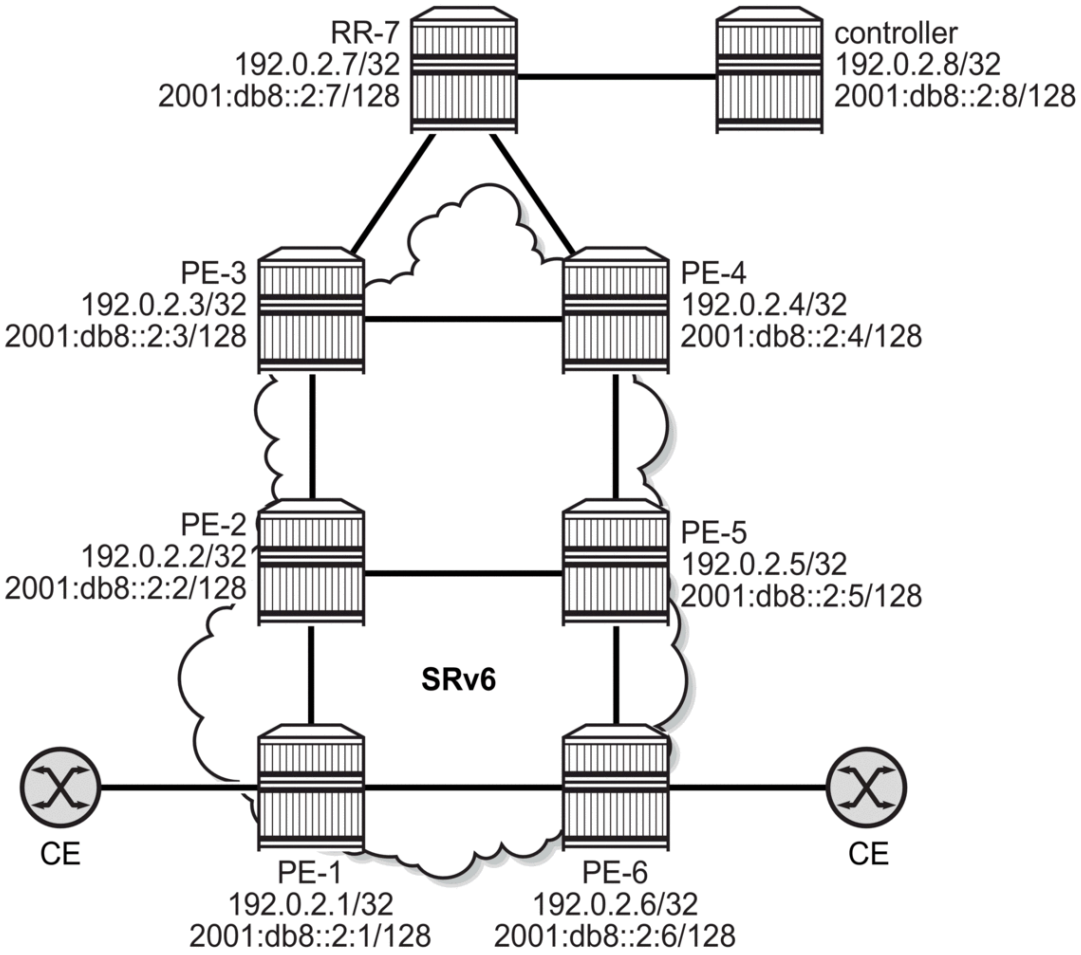
- **fallback-tunnel-to-route-table**

  This fallback command option first tries to resolve the route directly to a tunnel in the TTMv6. If none is found, the system falls back to the shortest path SRv6 resolution in the route table.

Configuration

Figure 68: Example topology with system IP addresses shows the example topology with six PEs, one route reflector (RR), and an external controller. In this example topology, the controller is an SR OS node that advertises non-local SRv6 policies using BGP. The controller is only used in two of the following examples; in the other examples, static SRv6 policies are configured on PE-1. The CEs are connected to PE-1 and PE-6 when Layer 2 services are configured.

Figure 68: Example topology with system IP addresses



The initial configuration includes:

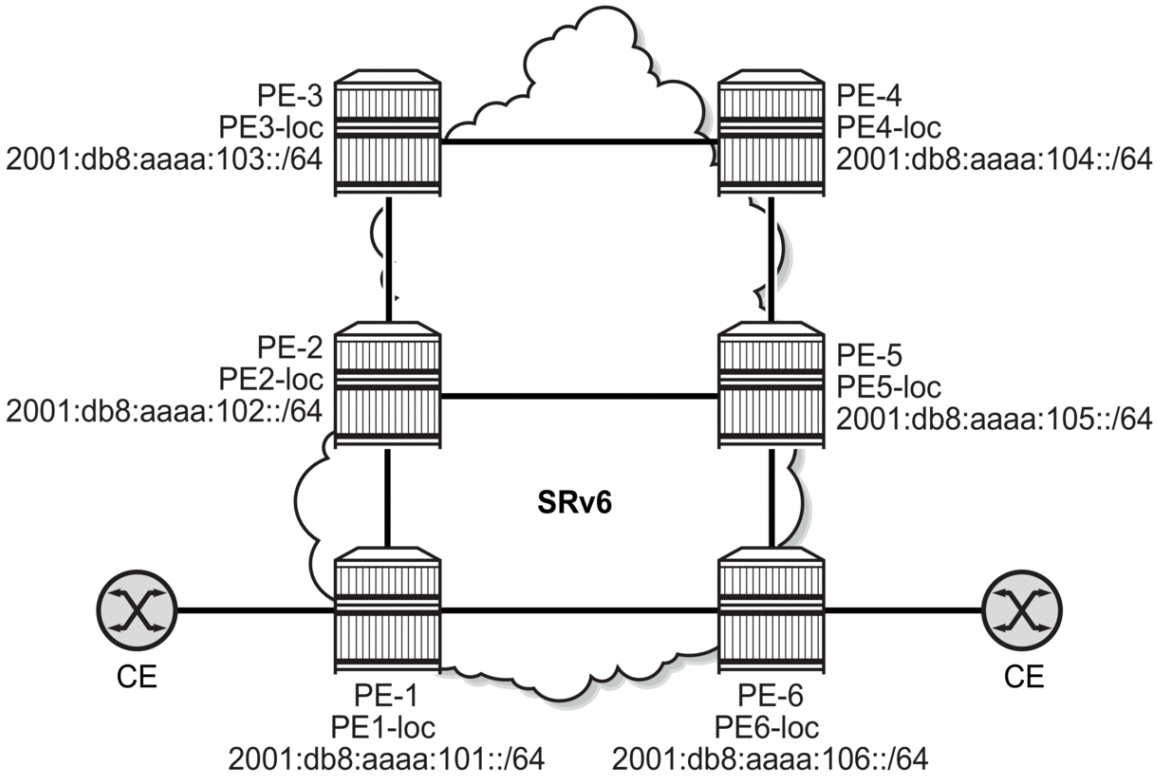
- cards, MDAs, ports
- router interfaces with IPv6 addresses – only the system interface is configured with both IPv4 and IPv6 addresses
- IS-IS on all router interfaces between the PEs and between PE-3, PE-4, and RR-7
- static routes between RR-7 and the controller
- SRv6 between the PEs, but not toward RR-7

SRv6 locators and BGP configuration

Classic SRv6 locators

Figure 69: Example topology with classic locator prefixes shows the classic SRv6 locators that are configured on the PEs.

Figure 69: Example topology with classic locator prefixes



39571

As an example, on PE-2, the classic locator "PE2-loc" is configured as follows:

```
# on PE-2:
configure {
    router "Base" {
        mpls-labels {
            reserved-label-block "sr-policy" {
                start-label 20000
                end-label 22999
            }
        }
        isis 0 {
            ---snip---
            segment-routing-v6 {
                admin-state enable
```

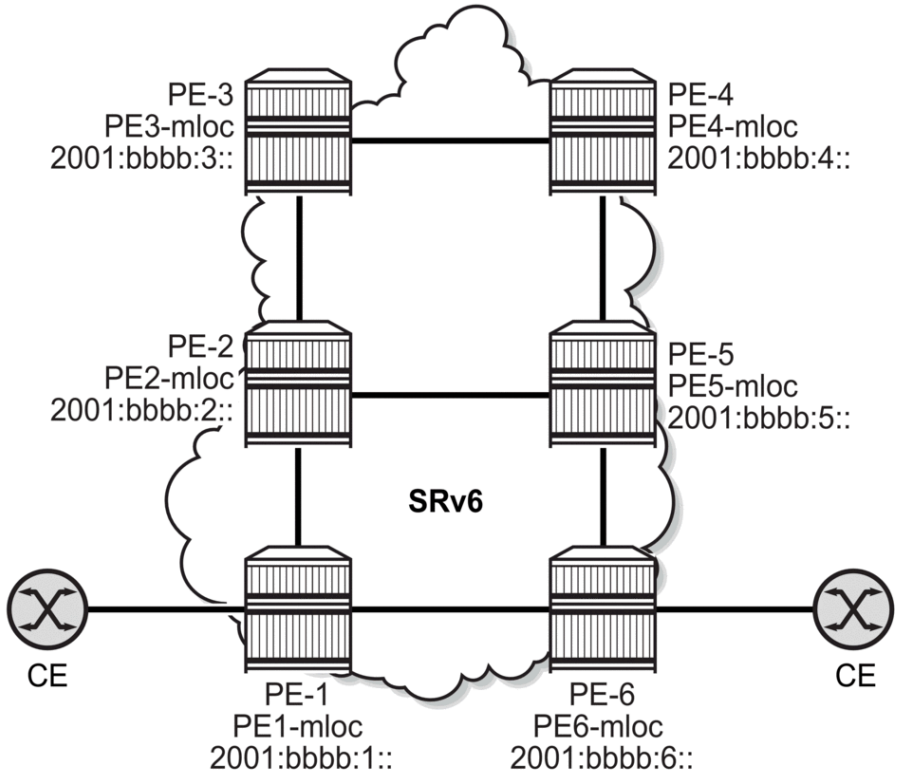
```
                locator "PE2-loc" {
                }
            }
        }
        segment-routing {
            segment-routing-v6 {
                origination-fpe [1]
                source-address 2001:db8::2:2
                locator "PE2-loc" {
                    admin-state enable
                    block-length 48
                    termination-fpe [2]
                    prefix {
                        ip-prefix 2001:db8:aaaa:102::/64
                    }
                    static-function {
                        max-entries 16
                        label-block "sr-policy"
                    }
                }
                base-routing-instance {
                    locator "PE2-loc" {
                        function {
                            end 1 {
                                srh-mode usp
                            }
                            end-x 2 {     # static function value for adj with PE-1
                                protection unprotected
                                interface-name "int-PE-2-PE-1"
                            }
                            end-x 3 {     # static function value for adj with PE-3
                                protection unprotected
                                interface-name "int-PE-2-PE-3"
                            }
                            end-x-auto-allocate psp protection unprotected { }
                        }
                    }
                }
```

The configuration on the other PEs is similar.

Micro-segment SRv6 locators

Besides the classic SRv6 locators, the PEs are also configured with micro-segment SRv6 locators. Figure 70: Example topology with micro-segment node SIDs shows the micro-segment node SIDs configured on the PEs.

Figure 70: Example topology with micro-segment node SIDs



On PE-2, the micro-segment locator "PE2-mloc" is configured as follows:

```
# on PE-2:
configure {
    router "Base" {
        mpls-labels {
            reserved-label-block "res-block1" {
                start-label 19000
                end-label 19999
            }
        }
        isis 0 {
            segment-routing-v6 {
                admin-state enable
                micro-segment-locator "PE2-mloc" {
                }
            }
        }
        segment-routing {
            segment-routing-v6 {
                origination-fpe [1]
                source-address 2001:db8::2:2
                micro-segment {
                    argument-length 16
                    block "ms-block1" {
                        admin-state enable
```

```
                    termination-fpe [2]
                    label-block "res-block1"
                    prefix {
                        ip-prefix 2001:bbbb::/32
                    }
                    static-function {
                        max-entries 16
                    }
                }
            }
            micro-segment-locator "PE2-mloc" {
                admin-state enable
                block "ms-block1"
                un {
                    srh-mode usp
                    value 2
                }
            }
            base-routing-instance {
                micro-segment-locator "PE2-mloc" {
                    function {
                        ua 1 {
                            protection unprotected
                            interface-name "int-PE-2-PE-1"
                        }
                        ua 2 {
                            protection unprotected
                            interface-name "int-PE-2-PE-3"
                        }
                        ua 3 {
                            protection unprotected
                            interface-name "int-PE-2-PE-5"
                        }
                        ua-auto-allocate psp protection unprotected { }
                    }
                }
            }
        }
```

The configuration on the other PEs is similar.

BGP configuration

The following examples are described in this section:

- EVPN VPWS
- EVPN VPLS
- EVPN IFL
- IP VPN

For the examples where EVPN VPWS and EVPN VPLS services are configured and static SRv6 policies are used, it is sufficient to configure BGP for the EVPN address family only, as follows:

```
# on PE-1, PE-2, PE-3, PE-4, PE-5, PE-6:
configure {
    router "Base" {
        autonomous-system 64500
        bgp {
            vpn-apply-export true
```

```
            vpn-apply-import true
            rapid-withdrawal true
            peer-ip-tracking true
            split-horizon true
            rapid-update {
                evpn true
            }
            group "internal" {
                peer-as 64500
                family {
                    evpn true
                }
            }
            neighbor "2001:db8::2:7" {
                group "internal"
            }
        }
```

For the EVPN VPWS and EVPN VPLS services where BGP-signaled SRv6 policies are used, BGP must also be enabled for the SRv6 policy address family, as follows:

```
# on PE-6:
configure {
    router "Base" {
        autonomous-system 64500
        bgp {
            vpn-apply-export true
            vpn-apply-import true
            rapid-withdrawal true
            peer-ip-tracking true
            split-horizon true
            rapid-update {
                evpn true
            }
            group "internal" {
                peer-as 64500
                family {
                    evpn true
                    sr-policy-ipv6 true
                }
            }
            neighbor "2001:db8::2:7" {
                group "internal"
            }
        }
```

RR-7 also has a BGP session with the controller. The BGP configuration on RR-7 is as follows:

```
# on RR-7:
configure {
    router "Base" {
        autonomous-system 64500
        bgp {
            vpn-apply-export true
            vpn-apply-import true
            rapid-withdrawal true
            peer-ip-tracking true
            split-horizon true
            rapid-update {
                evpn true
            }
            group "IBGP clients" {
```

```
                peer-as 64500
                family {
                    evpn true
                    sr-policy-ipv6 true
                }
                cluster {
                    cluster-id 192.0.2.7
                }
            }
            group "SRv6-policies" {
                peer-as 64500
                family {
                    ipv6 true
                    sr-policy-ipv6 true
                }
            }
            neighbor "2001:db8::2:1" {
                group "IBGP clients"
            }
            neighbor "2001:db8::2:2" {
                group "IBGP clients"
            }
            neighbor "2001:db8::2:3" {
                group "IBGP clients"
            }
            neighbor "2001:db8::2:4" {
                group "IBGP clients"
            }
            neighbor "2001:db8::2:5" {
                group "IBGP clients"
            }
            neighbor "2001:db8::2:6" {
                group "IBGP clients"
            }
            neighbor "2001:db8::2:8" {
                group "SRv6-policies"
            }
        }
```

The controller signals SRv6 policies using BGP. The **sr-policy-import true** command instructs BGP to import statically-configured non-local segment routing policies from the segment routing database into the BGP RIB so that they can be advertised, as originated routes, toward BGP peers that support the **sr-policy-ipv6** address family. The BGP configuration on the controller is as follows:

```
# on controller:
configure {
    router "Base" {
        autonomous-system 64500
        bgp {
            rapid-withdrawal true
            sr-policy-import true   # import static non-local SR policies into BGP RIB
            group "SR-policy" {
                peer-as 64500
                family {
                    ipv6 true
                    sr-policy-ipv6 true
                }
            }
            neighbor "2001:db8::2:7" {
                group "SR-policy"
            }
        }
```

For the EVPN IFL services, the EVPN address family must be enabled and the IPv6 next hops must be advertised for the EVPN address family. In the case that an external controller is used, the SRv6 policy address family must also be enabled, as follows:

```
# on PE-1, PE-6:
configure {
    router "Base" {
        autonomous-system 64500
        bgp {
            vpn-apply-export true
            vpn-apply-import true
            rapid-withdrawal true
            peer-ip-tracking true
            split-horizon true
            rapid-update {
                evpn true
            }
            group "internal" {
                peer-as 64500
                family {
                    evpn true
                    sr-policy-ipv6 true
                }
                advertise-ipv6-next-hops {
                    evpn true
                }
            }
            neighbor "2001:db8::2:7" {
                group "internal"
            }
        }
```

For the IP VPN services, BGP must be configured for the VPN IPv4 and VPN IPv6 address families. If an external controller is used, the SRv6 policy address family must also be enabled. Extended next hop encoding is enabled for IPv4 and VPN IPv4; IPv6 next hops must be advertised for IPv4, VPN IPv4, and VPN IPv6. The BGP configuration on PE-1 and PE-6 is as follows:

```
# on PE-1, PE-6:
configure {
    router "Base" {
        bgp {
            vpn-apply-export true
            vpn-apply-import true
            rapid-withdrawal true
            peer-ip-tracking true
            split-horizon true
            rapid-update {
            }
            group "internal" {
                peer-as 64500
                family {
                    vpn-ipv4 true
                    vpn-ipv6 true
                    sr-policy-ipv6 true
                }
                extended-nh-encoding {
                    vpn-ipv4 true
                    ipv4 true
                }
                advertise-ipv6-next-hops {
                    vpn-ipv6 true
                    vpn-ipv4 true
```

```
                ipv4 true
            }
        }
        neighbor "2001:db8::2:7" {
            group "internal"
        }
    }
```

EVPN VPWS

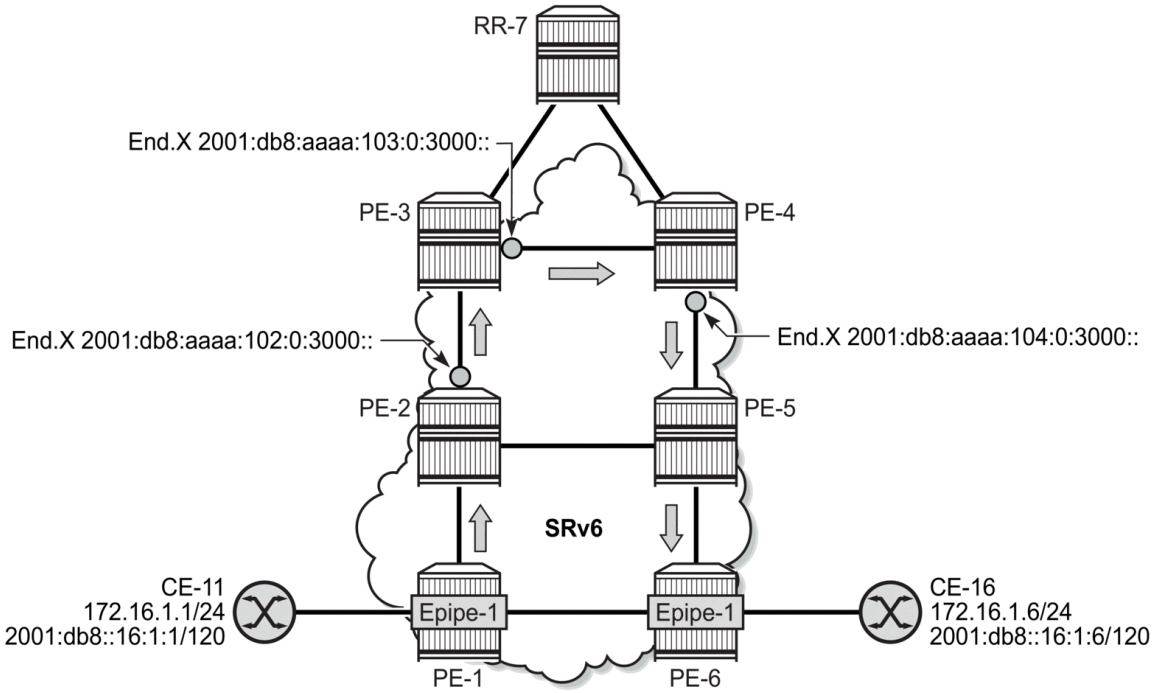
The following examples are described in this section:

- EVPN VPWS with static SRv6 policy on head-end PE-1
- EVPN VPWS with BGP-signaled SRv6 policy on head-end PE-6

EVPN VPWS with static SRv6 policy on head-end PE-1

Figure 71: EVPN VPWS using SRv6 policy with color 100 from PE-1 to PE-6 shows the segments in the segment list of the static SRv6 policy on head-end PE-1. The SIDs are the End.X SID for the interface "int-PE-2-PE-3" on PE-2, the End.X SID for the interface "int-PE-3-PE-4" on PE-3, and the End.X SID for the interface "int-PE-4-PE-5" on PE-4.

Figure 71: EVPN VPWS using SRv6 policy with color 100 from PE-1 to PE-6



Static SRv6 policy on PE-1

SRv6 policies are configured with the following command:

```
[ex:/configure router "Base" segment-routing sr-policies static-policy "color-100-PE-1-PE-6"]
A:admin@PE-1# ?

 admin-state           - Administrative state of segment routing static policy
 apply-groups          - Apply a configuration group at this level
 apply-groups-exclude  - Exclude a configuration group at this level
 binding-sid           - Segment ID that opaquely represents an SR policy to upstream routers
 color                 - Traffic flows to be steered by this policy
 distinguisher         - Unique value for a policy
 endpoint              - Destination of the source-routed path
 head-end              - Head end address for this static policy
 maintenance-policy    - Policy name
 preference            - Preference value of this static policy
 segment-list          + Enter the segment-list list instance
 segment-routing-v6    + Enable the segment-routing-v6 context
 type                  - Static policy type

[ex:/configure router "Base" segment-routing sr-policies static-policy "color-100-PE-1-PE-6"]
A:admin@PE-1# type ?

 type <keyword>
 <keyword> - (sr-mpls|srv6)
 Default   - sr-mpls

    Static policy type
```

The type of SR policy is SRv6 and the endpoint must be an IPv6 address. In the following example, the static SRv6 policy is configured on head-end PE-1 itself; therefore, the head-end is set to **local** and the binding SID cannot be configured as an IPv6 address for a local SRv6 policy:

```
*[ex:/configure router "Base" segment-routing sr-policies static-policy "color-150-PE-1-PE-6"
 segment-routing-v6 binding-sid 1]
A:admin@PE-1# ip-address 2001:db8:aaaa:101:0:5000::

*[ex:/configure router "Base" segment-routing sr-policies static-policy "color-150-PE-1-PE-6"
 segment-routing-v6 binding-sid 1]
A:admin@PE-1# commit
MINOR: SRDB #12: configure router "Base" segment-routing sr-policies static-policy "color-150-
PE-1-PE-6" segment-routing-v6 binding-sid 1 ip-address - Inconsistent Value error - head-end
 must be remote ip-address when ip-address is configured - configure router "Base" segment-
routing sr-policies static-policy "color-150-PE-1-PE-6" head-end
```

The only way to configure this binding SID 2001:db8:aaaa:101:0:5000:: in the static SRv6 policy on the local head-end is by specifying a function value from the static range, in this example **function-value 5**, as follows:

```
# on PE-1:
configure {
    router "Base" {
        segment-routing {
            sr-policies {
                static-policy "color-150-PE-1-PE-6" {
---snip---
                    segment-routing-v6 {
                        binding-sid 1 {
```

```
                            locator {
                                locator-name "PE1-loc"
                                function end-b6-encaps-red
                                function-value 5  # SID 2001:db8:aaaa:101:0:5000::
                            }
---snip---
```

When no function value is specified, the binding SID is automatically allocated from the dynamic range using the SRv6 locator prefix "PE1-loc" associated with the Base router.

When a classic SRv6 locator is used, the only supported binding SID behavior is End.B6.Encaps.Red, which is End.B6.Encaps with reduced Segment Routing Header (SRH). The reduced SRH does not contain the first SID. This first SID is only placed in the destination address of the pushed IPv6 header. The End.B6.Encaps endpoint behavior is bound to an SRv6 policy with encapsulation and the binding SID is instantiated by SRv6, as follows:

```
[ex:/configure router "Base" segment-routing sr-policies static-policy "color-100-PE-1-PE-6"
 segment-routing-v6 binding-sid 1 locator]
A:admin@PE-1# function ?

 function <keyword>
 <keyword> - end-b6-encaps-red

 'function' is: mandatory

    Behavior of the local SRv6 regular binding SID
```

In a similar way, when micro-segment SRv6 locators are used, the keyword to configure reduced encapsulation for the micro-binding SID is **end-b6-encaps-red-next-csid**; see further.

The following static SRv6 policy is configured on head-end PE-1:

```
# on head-end PE-1:
configure exclusive
    router "Base" {
        segment-routing {
            sr-policies {
                admin-state enable          # Enable SR policies context
                static-policy "color-100-PE-1-PE-6" {
                    admin-state enable      # Enable static policy
                    color 100
                    endpoint 2001:db8::2:6
                    head-end local
                    type srv6
                    segment-routing-v6 {
                        binding-sid 1 {
                            locator {
                                locator-name "PE1-loc"
                                function end-b6-encaps-red
                            }
                        }
                    }
                    segment-list 1 {
                        admin-state enable      # Enable segment list 1
                        segment 1 {
                            srv6-sid 2001:db8:aaaa:102:0:3000:: # End.X int-PE-2-PE-3
                        }
                        segment 2 {
                            srv6-sid 2001:db8:aaaa:103:0:3000:: # End.X int-PE-3-PE-4
                        }
                        segment 3 {
                            srv6-sid 2001:db8:aaaa:104:0:3000:: # End.X int-PE-4-PE-5
```

```
                    }
                }
            }
```

The SRv6 SIDs in segment list 1 are the following adjacency SIDs:

- 2001:db8:aaaa:102:0:3000:: for the End.X of interface "int-PE-2-PE-3" on PE-2
- 2001:db8:aaaa:103:0:3000:: for the End.X of interface "int-PE-3-PE-4" on PE-3
- 2001:db8:aaaa:104:0:3000:: for the End.X of interface "int-PE-4-PE-5" on PE-4

The following list of adjacency SIDs on PE-2 shows that SID 2001:db8:aaaa:102:0:3000:: corresponds to the End.X SID of "int-PE-2-PE-3":

```
[/]
A:admin@PE-2# show router segment-routing-v6 base-routing-instance end-x

===============================================================================
Segment Routing v6 Base Routing Instance
===============================================================================
Locator
 Type          Function  SID                                     Status/InstId
  SRH-mode     Oper Func Interface                               Protection
-------------------------------------------------------------------------------
PE2-loc
 End.X         2         2001:db8:aaaa:102:0:2000::              ok
  PSP          2         int-PE-2-PE-1                           Unprotected
 End.X         3         2001:db8:aaaa:102:0:3000::              ok
  PSP          3         int-PE-2-PE-3                           Unprotected
-------------------------------------------------------------------------------
 Auto-allocated End.X:
-------------------------------------------------------------------------------
 End.X         17        2001:db8:aaaa:102:1:1000::              0
  PSP                    int-PE-2-PE-5                           Unprotected
  ISIS Level: L2 Mac Address: 02:1e:01:01:00:15 Nbr Sys Id: 1920.0000.2005
-------------------------------------------------------------------------------
===============================================================================
Legend: * - System allocated
```

The following shows the summary of the SR policies with only one static SRv6 policy configured and active on PE-1:

```
[/]
A:admin@PE-1# show router segment-routing sr-policies summary

===============================================================================
SR-Policies Summary
===============================================================================
Admin Status          : Up
Ingress Stats         : N/A
Egress Stats          : N/A
Resv Label Blk Name:
TTM Preference        : 14
SR-MPLS BSID Allocated: 0              SRV6 BSID Allocated   : 1
Static Local Policies : 1              Active Static Lcl Pol : 1
Static Non Local Pol  : 0
BGP Policies          : 0              Active BGP Policies   : 0
===============================================================================
```

The following command on PE-1 shows the static SR policy with color 100, head-end 0.0.0.0 for local, and endpoint 2001:db8::2:6. The binding SID 2001:db8:aaaa:101:1:2000:: is automatically allocated using the locator prefix "PE1-loc" associated with the Base router.

```
[/]
A:admin@PE-1# show router segment-routing sr-policies static

===============================================================================
SR-Policies Path
===============================================================================
-------------------------------------------------------------------------------
Type            : srv6
Active          : Yes                   Owner           : static
Operational     : Yes
Color           : 100
Head            : 0.0.0.0               Endpoint Addr   : 2001:db8::2:6
RD              : 0                     Preference      : 100
SRv6 BSID 1     : 2001:db8:aaaa:101:1:2000::
TunnelId        : 917506                Age             : 1084
Origin ASN      : 0                     Origin          : 0.0.0.0
NumReEval       : 0                     LastReEvalReason: none
NumActPathChange: 0                     Last Change     : 06/13/2024 13:08:24
Maintenance Plcy:
Ret Path BFD SID:

Path Segment Lists:
Segment-List    : 1                     Weight          : 1
Num Segments    : 3                     Last Change     : 06/13/2024 12:57:02
  1 SRv6 SID  : 2001:db8:aaaa:102:0:3000::               State : resolved-up
  2 SRv6 SID  : 2001:db8:aaaa:103:0:3000::               State : N/A
  3 SRv6 SID  : 2001:db8:aaaa:104:0:3000::               State : N/A

===============================================================================
```

The SRv6 policy tunnel on PE-1 with tunnel ID 917506 and preference 14 has color 100 and the SRv6 SIDs are 2001:db8:aaaa:102:0:3000::, 2001:db8:aaaa:103:0:3000::, and 2001:db8:aaaa:104:0:3000::, as follows:

```
[/]
A:admin@PE-1# show router tunnel-table ipv6 protocol srv6-policy detail

===============================================================================
Tunnel Table (Router: Base)
===============================================================================
Destination      : 2001:db8::2:6/128
NextHop          : fpe_1.a
NextHop Weight   : 1
Tunnel Flags     : has-color
Age              : 00h18m04s            Color            : 100
CBF Classes      : (Not Specified)
Owner            : srv6-pol             Encap            : SRV6
Tunnel ID        : 917506               Preference       : 14
Tunnel SRV6 SID  : 2001:db8:aaaa:102:0: Tunnel Metric    : 0
                   3000::
                 : 2001:db8:aaaa:103:0:3000::
                 : 2001:db8:aaaa:104:0:3000::
Tunnel MTU       :  -                   Max Label Stack  : 3
-------------------------------------------------------------------------------
Number of tunnel-table entries          : 1
Number of tunnel-table entries with LFA : 0
===============================================================================
```

The FP tunnel table on PE-1 shows the SRv6 policy, including the SID list, as follows:

```
[/]
A:admin@PE-1# show router fp-tunnel-table 1 ipv6 protocol srv6-policy

===============================================================================
IPv6 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                Protocol         Tunnel-ID
  Lbl/SID
    NextHop                                                 Intf/Tunnel
  Lbl/SID (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
2001:db8::2:6/128                          SRV6-Policy       -
  2001:db8:aaaa:104:0:3000::
  2001:db8:aaaa:103:0:3000::
  2001:db8:aaaa:102:0:3000::
    0.140.1.1                                              pxc-1.b:1
-------------------------------------------------------------------------------
Total Entries : 1
-------------------------------------------------------------------------------
===============================================================================
```

EVPN VPWS Epipe-1 on PE-1 and PE-6

On PE-1, EVPN VPWS "Epipe-1" uses SRv6 transport and is configured with **resolution fallback-tunnel-to-route-table**, as follows:

```
# on PE-1:
configure {
    service {
        epipe "Epipe-1" {
            admin-state enable
            service-id 1
            customer "1"
            segment-routing-v6 1 {
                locator "PE1-loc" {
                    function {
                        end-dx2 {
                        }
                    }
                }
            }
            bgp 1 {
                route-target {
                    export "target:64500:1"
                    import "target:64500:1"
                }
            }
            sap 1/1/c10/1:1 {
                description "SAP to CE-11"
            }
            bgp-evpn {
                evi 1
                local-attachment-circuit "PE1" {
```

```
                    eth-tag 1
                }
                remote-attachment-circuit "PE6" {
                    eth-tag 6
                }
                segment-routing-v6 1 {
                    admin-state enable
                    source-address 2001:db8::2:1
                    resolution fallback-tunnel-to-route-table
                    srv6 {
                        instance 1
                        default-locator "PE1-loc"
                    }
                    route-next-hop {
                        ip-address 2001:db8::2:1
                    }
                }
            }
        }
```

The following command shows the BGP EVPN SRv6 information for Epipe-1 on PE-1:

```
[/]
A:admin@PE-1# show service id "Epipe-1" bgp-evpn segment-routing-v6

===============================================================================
BGP EVPN Segment Routing v6 Information
===============================================================================
Admin State             : Enabled                Bgp Instance  : 1
Srv6 Instance           : 1
Default Locator         : PE1-loc

Oper Group              : (none)
Default Route Tag       : 0x0
Source Address          : 2001:db8::2:1
ECMP                    : 1
Force Vlan VC Fwd       : Disabled
Next Hop Type           : explicit
Next Hop Address        : 2001:db8::2:1
Evi 3-byte Auto-RT      : disabled
Route Resolution        : fallback-tunnel-to-route-table
Force QinQ VC Fwd       : none
MH Mode                 : network
Domain-Id               : None
===============================================================================
```

The static SRv6 policy is identified by the color 100 (color:00:100), head-end 192.0.2.1 (system IPv4 address of PE-1), and endpoint 2001:db8::2:6 (system IPv6 address of PE-6). PE-6 exports color 100 in an export policy to match the color in the SRv6 policy on PE-1, as follows:

```
#  on PE-6
configure exclusive
    policy-options {
        community "color-100" {
            member "color:00:100" { }
        }
        community "vsi-1" {
            member "target:64500:1" { }
        }
        policy-statement "epipe-1-export-c100" {
            default-action {
                action-type accept
```

```
                community {
                    add ["vsi-1" "color-100"]
                }
            }
        }
    }
    service {
        epipe "Epipe-1" {
            admin-state enable
            service-id 1
            customer "1"
            segment-routing-v6 1 {
                locator "PE6-loc" {
                    function {
                        end-dx2 {
                        }
                    }
                }
            }
            bgp 1 {
                vsi-export ["epipe-1-export-c100"]
                route-target {
                    import "target:64500:1"
                }
            }
            sap 1/1/c10/1:1 {
                description "SAP to CE-16"
            }
            bgp-evpn {
                evi 1
                local-attachment-circuit "PE6" {
                    eth-tag 6
                }
                remote-attachment-circuit "PE1" {
                    eth-tag 1
                }
                segment-routing-v6 1 {
                    admin-state enable
                    source-address 2001:db8::2:6
                    srv6 {
                        instance 1
                        default-locator "PE6-loc"
                    }
                    route-next-hop {
                        ip-address 2001:db8::2:6
                    }
                }
            }
        }
```

The following EVPN auto-discovery route received on PE-1 includes the community color:00:100 (color 100):

```
[/]
A:admin@PE-1# show router bgp routes evpn auto-disc rd 192.0.2.6:1 detail
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================
 Legend -
 Status codes  : u - used, s - suppressed, h - history, d - decayed, * - valid
                 l - leaked, x - stale, > - best, b - backup, p - purge
 Origin codes  : i - IGP, e - EGP, ? - incomplete
```

```
===============================================================================
BGP EVPN Auto-Disc Routes
===============================================================================
Original Attributes

Network        : n/a
Nexthop        : 2001:db8::2:6
Path Id        : None
From           : 2001:db8::2:7
Res. Nexthop   : fe80::24:1ff:fe01:1
Local Pref.    : 100                    Interface Name : int-PE-1-PE-6
Aggregator AS  : None                   Aggregator     : None
Atomic Aggr.   : Not Atomic             MED            : None
AIGP Metric    : None                   IGP Cost       : 10
Connector      : None
Community      : target:64500:1 color:00:100
                 l2-attribute:MTU: 1514 F: 0 C: 0 P: 0 B: 0
Cluster        : 192.0.2.7
Originator Id  : 192.0.2.6              Peer Router Id : 192.0.2.7
Origin         : IGP
Flags          : Used Valid Best
Route Source   : Internal
AS-Path        : No As-Path
EVPN type      : AUTO-DISC
ESI            : ESI-0
Tag            : 6
Route Dist.    : 192.0.2.6:1
MPLS Label     : 504288
Route Tag      : 0
Neighbor-AS    : n/a
DB Orig Val    : N/A                    Final Orig Val : N/A
Source Class   : 0                      Dest Class     : 0
Add Paths Send : Default
Last Modified  : 00h00m47s
SRv6 TLV Type  : SRv6 L2 Service TLV (6)
SRv6 SubTLV    : SRv6 SID Information (1)
Sid            : 2001:db8:aaaa:106::
Full Sid       : 2001:db8:aaaa:106:7b1e::
Behavior       : End.DX2 (21)
SRv6 SubSubTLV : SRv6 SID Structure (1)
Loc-Block-Len  : 48                     Loc-Node-Len   : 16
Func-Len       : 20                     Arg-Len        : 0
Tpose-Len      : 20                     Tpose-offset   : 64
---snip---
```

On PE-1, the BGP next hop for EVPN routes to 2001:db8::2:6 is resolved using the SRv6 policy with tunnel ID 917506, as follows:

```
[/]
A:admin@PE-1# show router bgp next-hop evpn 2001:db8::2:6 detail
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================

===============================================================================
BGP VPN Next Hop
===============================================================================
-------------------------------------------------------------------------------
VPN Next Hop          : 2001:db8::2:6
Autobind              : gre/rtm srv6-policy
Labels                : --
User-labels           : 1
Admin-tag-policy      : --
```

```
Strict-tunnel-tagging : N
Color                 : 100
UPA Trigger Next Hop  : --
Locator               : 2001:db8:aaaa:106::/64
Created               : 00h01m16s
Last-modified         : 00h01m16s
-------------------------------------------------------------------------------
Resolving Prefix : 2001:db8::2:6/128
Preference       : 14                   Metric           : 0
Reference Count  : 1                    Owner            : SRV6-POLICY
Fib Programmed   : Y
Resolved Next Hop: 0.140.1.1
Egress Label     : n/a                  TunnelId         : 917506
Locator State    : Resolved
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
Next Hops : 1
===============================================================================
```

Resolution fallback

While PE-6 exports color 100, the corresponding SRv6 policy tunnel on PE-1 is available, as follows:

```
[/]
A:admin@PE-1# show router tunnel-table ipv6 protocol srv6-policy

===============================================================================
IPv6 Tunnel Table (Router: Base)
===============================================================================
Destination                                 Owner     Encap TunnelId  Pref
Nexthop                                     Color           Metric
-------------------------------------------------------------------------------
2001:db8::2:6/128                           srv6-pol  SRV6  917506    14
  fpe_1.a                                     100             0
-------------------------------------------------------------------------------
Flags: B = BGP or MPLS backup hop available
       L = Loop-Free Alternate (LFA) hop available
       E = Inactive best-external BGP route
       k = RIB-API or Forwarding Policy backup hop
===============================================================================
```

```
[/]
A:admin@PE-1# show router route-table ipv6 protocol srv6-policy

===============================================================================
IPv6 Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                          Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
2001:db8:aaaa:101:1:2000::/128              Local   SRV6-Pol* 00h17m06s  14
       2001:db8::2:6 (tunneled:SRV6-Policy:917506)                  1
-------------------------------------------------------------------------------
No. of Routes: 1
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
===============================================================================
* indicates that the corresponding row element may have been truncated.
```

When PE-6 does not export color 100, there is no match with the SRv6 policy on PE-1, and PE-1 uses the route table instead of the tunnel table. The export policy on PE-6 is removed in Epipe-1, as follows:

```
# on PE-6:
configure {
    service {
        epipe "Epipe-1" {
            bgp 1 {
                delete vsi-export "epipe-1-export-c100"
                route-target {
                    export "target:64500:1"
                    import "target:64500:1"
                }
            }
```

On PE-1, the BGP next hop for the EVPN routes to 2001:db8::2:6 is not resolved to an SRv6 policy tunnel anymore; the route table is used instead, as follows:

```
[/]
A:admin@PE-1# show router bgp next-hop 2001:db8::2:6 evpn service-id 1
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================

===============================================================================
BGP VPN Next Hop
===============================================================================
VPN Next Hop                                                    Owner
   Autobind                                          FibProg  Reason
   Labels (User-labels)                              FlexAlgo Metric
   Admin-tag-policy (strict-tunnel-tagging)                   Last Mod.
-------------------------------------------------------------------------------
2001:db8::2:6                                                   GRE/RTM
   gre/rtm srv6-policy                               Y
   -- (1)                                            --       10
   -- (N)                                                     00h00m24s
-------------------------------------------------------------------------------
Next Hops : 1
===============================================================================

===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================
```

When the SRv6 resolution is reconfigured to the tunnel table only instead of the fallback, the next hop cannot be resolved anymore, as follows:

```
# on PE-1:
configure exclusive
    service {
        epipe "Epipe-1" {
            bgp-evpn {
                segment-routing-v6 1 {
                    resolution tunnel-table
```

The BGP next hop cannot be resolved and the reason is a color mismatch, as follows:

```
*[/]
A:admin@PE-1# show router bgp next-hop 2001:db8::2:6 evpn service-id 1
===============================================================================
```

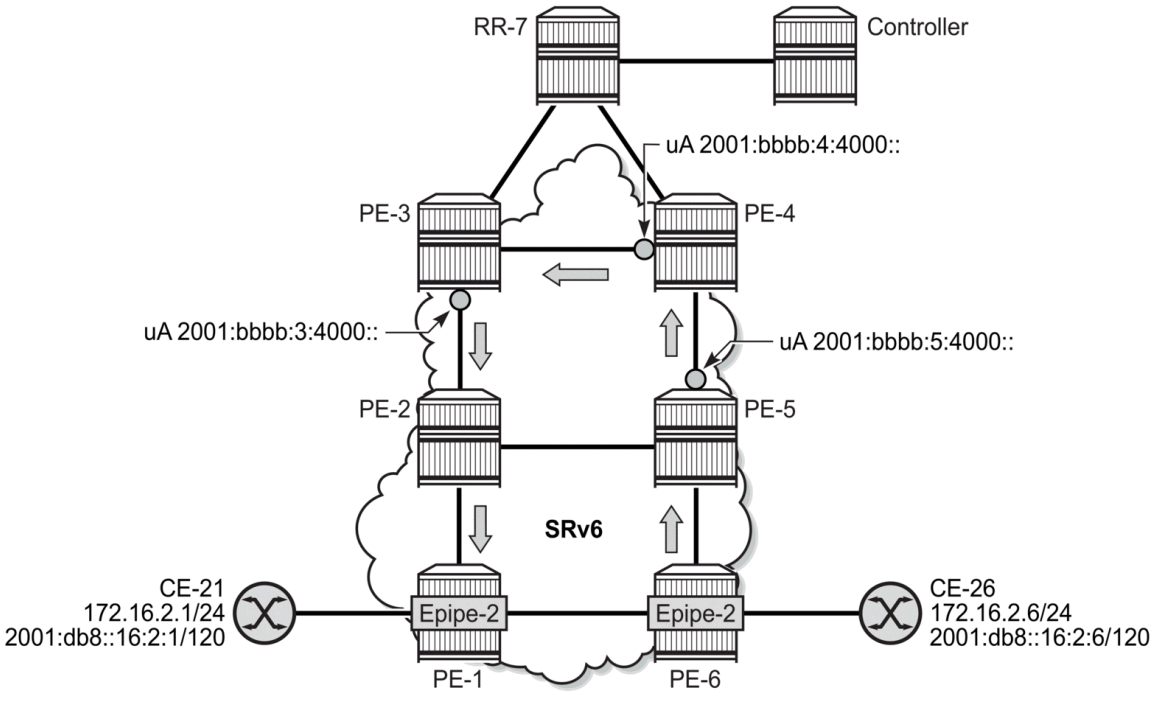
```
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================

===============================================================================
BGP VPN Next Hop
===============================================================================
VPN Next Hop                                                    Owner
   Autobind                                          FibProg    Reason
   Labels (User-labels)                              FlexAlgo   Metric
   Admin-tag-policy (strict-tunnel-tagging)                     Last Mod.
-------------------------------------------------------------------------------
2001:db8::2:6                                                   --
   srv6-policy                                       N          ColorMismatch
   -- (1)                                            --
   -- (N)                                                       00h00m05s
-------------------------------------------------------------------------------
Next Hops : 1
===============================================================================
```

When PE-6 exports color 100 again, the BGP next hop can be resolved using the SRv6 policy.

EVPN VPWS with BGP-signaled SRv6 policy on head-end PE-6

Figure 72: EVPN VPWS using SRv6 policy with color 200 from PE-6 to PE-1 shows the segments in the segment list of the BGP-signaled SRv6 policy on head-end PE-6. This SRv6 policy is configured on the controller, imported in the BGP RIB, and advertised using BGP. The SIDs in the segment list of the SRv6 policy are the micro-segment adjacency SID (uA SID) for the interface "int-PE-5-PE-4" on PE-5, the uA SID for the interface "int-PE-4-PE-3" on PE-4, and the uA SID for the interface "int-PE-3-PE-2" on PE-3. These micro-SIDs are configured with behavior and structure information. This information must only be configured when the SRv6 SID is a micro-SID. In fact, the allowed behaviors which describe the SRv6 SID are only micro-segment behaviors. By configuring this information, the user indirectly provides the head-end with necessary information to apply a compression algorithm to the received micro-SIDs. The algorithm is designed to reduce the Segment Routing Header (SRH) size. Details on what behavior and structure values to configure based on the SRv6 SID, and on how the algorithm processes these are described in the 7750 SR and 7950 XRS Segment Routing and PCE User Guide.

Figure 72: EVPN VPWS using SRv6 policy with color 200 from PE-6 to PE-1



Static SRv6 policy on the controller

The following SRv6 policy with micro-segment SIDs is configured on the controller:

```
# on controller:
configure {
    router "Base" {
        segment-routing {
            sr-policies {
                admin-state enable
                static-policy "color-200-PE-6-PE-1-u" {
                    admin-state enable
                    color 200
                    endpoint 2001:db8::2:1      # IPv6 system address of PE-1
                    head-end 192.0.2.6          # IPv4 system address of PE-6
                    distinguisher 200006001     # unique value
                    type srv6
                    segment-routing-v6 {
                        binding-sid 1 {
                            ip-address 2001:bbbb:6:4005::   # available address
                        }
                    }
                    segment-list 1 {
                        admin-state enable
                        segment 1 {
                            srv6-sid 2001:bbbb:5:4000:: # uA for "int-PE-5-PE-4"
                            behavior-and-structure {
                                behavior end-x-next-csid
```

```
                            block-length 32
                            node-length 16
                            function-length 16
                        }
                    }
                    segment 2 {
                        srv6-sid 2001:bbbb:4:4000:: # uA for "int-PE-4-PE-3"
                        behavior-and-structure {
                            behavior end-x-next-csid
                            block-length 32
                            node-length 16
                            function-length 16
                        }
                    }
                    segment 3 {
                        srv6-sid 2001:bbbb:3:4000:: # uA for "int-PE-3-PE-2"
                        behavior-and-structure {
                            behavior end-x-next-csid
                            block-length 32
                            node-length 16
                            function-length 16
                        }
                    }
                }
            }
```

The distinguisher must be unique and is used as a differentiator. If the head-end PE receives two policies with the same color and endpoint, the path with the lowest distinguisher value is chosen.

The head-end is the IPv4 system address of PE-6; the endpoint is the IPv6 system address of PE-1.

The binding SID is an available IPv6 address from the static function entries on head-end PE-6.

Note:
At the configuration point, the micro-binding SID must be configured in the <block><uN><uB6> format, because the head-end only accepts this format. The head-end then programs two route entries for each local SID: <block><uN><uB6>::/block-length+32 and <block><uB6>::/block-length+16, both pointing to the SRv6 policy.

The controller advertises this SRv6 policy using BGP. The head-end PE-6 must have the SR policies context enabled, as follows:

```
# on head-end PE-6:
configure {
    router "Base" {
        segment-routing {
            sr-policies {
                admin-state enable     # enable context for (BGP-signaled) SRv6 policies
            }
```

EVPN VPWS Epipe-2 on PE-1 and PE-6

PE-1 exports color 200 (color:00:200) in Epipe-2, as follows:

```
# on PE-1:
configure {
    policy-options {
        community "color-200" {
            member "color:00:200" { }
```

```
        }
        community "vsi-2" {
            member "target:64500:2" { }
        }
        policy-statement "epipe-2-export-c200" {
            default-action {
                action-type accept
                community {
                    add ["vsi-2" "color-200"]
                }
            }
        }
    }
    service {
        epipe "Epipe-2" {
            admin-state enable
            service-id 2
            customer "1"
            segment-routing-v6 1 {
                micro-segment-locator "PE1-mloc" {
                    function {
                        udx2 {
                            value 5
                        }
                    }
                }
            }
            bgp 1 {
                vsi-export ["epipe-2-export-c200"]
                route-target {
                    import "target:64500:2"
                }
            }
            sap 1/1/c10/1:2 {
                description "SAP to CE-21"
            }
            bgp-evpn {
                evi 2
                local-attachment-circuit "PE1" {
                    eth-tag 1
                }
                remote-attachment-circuit "PE6" {
                    eth-tag 6
                }
                segment-routing-v6 1 {
                    admin-state enable
                    source-address 2001:db8::2:1
                    srv6 {
                        instance 1
                        default-locator "PE1-mloc"
                    }
                    route-next-hop {
                        ip-address 2001:db8::2:1
                    }
                }
            }
        }
```

On PE-6, Epipe-2 is configured as follows:

```
# on head-end PE-6:
configure {
    service {
```

```
        epipe "Epipe-2" {
            admin-state enable
            service-id 2
            customer "1"
            segment-routing-v6 1 {
                micro-segment-locator "PE6-mloc" {
                    function {
                        udx2 {
                            value 5     # static value
                        }
                    }
                }
            }
            bgp 1 {
                route-target {
                    export "target:64500:2"
                    import "target:64500:2"
                }
            }
            sap 1/1/c10/1:2 {
                description "SAP to CE-26"
            }
            bgp-evpn {
                evi 2
                local-attachment-circuit "PE6" {
                    eth-tag 6
                }
                remote-attachment-circuit "PE1" {
                    eth-tag 1
                }
                segment-routing-v6 1 {
                    admin-state enable
                    source-address 2001:db8::2:6
                    resolution fallback-tunnel-to-route-table
                    srv6 {
                        instance 1
                        default-locator "PE6-mloc"
                    }
                    route-next-hop {
                        ip-address 2001:db8::2:6
                    }
                }
            }
        }
```

Verification

Head-end PE-6 receives one BGP-signaled SRv6 policy:

```
[/]
A:admin@PE-6# show router bgp summary all

===============================================================================
BGP Summary
===============================================================================
Legend : D - Dynamic Neighbor
===============================================================================
Neighbor
Description
ServiceId          AS PktRcvd InQ  Up/Down   State|Rcv/Act/Sent (Addr Family)
                      PktSent OutQ
```

```
-------------------------------------------------------------------------------
2001:db8::2:7
Def. Inst       64500      22    0 00h07m10s 2/2/2 (Evpn)
                           22    0           1/1/0 (SrPolicyIPv6)

-------------------------------------------------------------------------------
```

The following command shows the BGP-signaled SRv6 policy on head-end PE-6. The binding SID value is 2001:bbbb:4005:: and is derived from the binding SID 2001:bbbb:6:4005:: that was configured in the static SRv6 policy on the controller.

```
[/]
A:admin@PE-6# show router segment-routing sr-policies bgp

===============================================================================
SR-Policies Path
===============================================================================
-------------------------------------------------------------------------------
Type            : srv6
Active          : Yes                   Owner           : bgp
Operational     : Yes
Color           : 200
Head            : 0.0.0.0               Endpoint Addr   : 2001:db8::2:1
RD              : 200006001             Preference      : 100
SRv6 BSID 1     : 2001:bbbb:4005::
TunnelId        : 917506                Age             : 99
Origin ASN      : 64500                 Origin          : 2001:db8::2:8
NumReEval       : 0                     LastReEvalReason: none
NumActPathChange: 0                     Last Change     : 06/13/2024 13:37:00
Maintenance Plcy:
Ret Path BFD SID:

Path Segment Lists:
Segment-List    : 1                     Weight          : 1
Num Segments    : 3                     Last Change     : 06/13/2024 12:58:00
  1 SRv6 SID  : 2001:bbbb:5:4000::                       State : resolved-up
    Behavior & Structure
     Behavior   : 52                    Block Length    : 32
     Node Length: 16                    Function Length : 16
  2 SRv6 SID  : 2001:bbbb:4:4000::                       State : N/A
    Behavior & Structure
     Behavior   : 52                    Block Length    : 32
     Node Length: 16                    Function Length : 16
  3 SRv6 SID  : 2001:bbbb:3:4000::                       State : N/A
    Behavior & Structure
     Behavior   : 52                    Block Length    : 32
     Node Length: 16                    Function Length : 16

===============================================================================
```

On PE-6, the SID 2001:bbbb:6:4005:: has the value of the binding SID that was configured in the static SRv6 policy on the controller, as follows:

```
[/]
A:admin@PE-6# show router segment-routing-v6 micro-segment-local-sid

===============================================================================
Micro Segment Routing v6 Local SIDs
===============================================================================
SID                                             Type           Function
  Micro Segment Locator
  Context
```

```
-------------------------------------------------------------------------------
2001:bbbb:6::                                   uN             6
  PE6-mloc
  None
2001:bbbb:6:4000::                              uA             16384
  PE6-mloc
  Base
2001:bbbb:6:4004::                              uDX2           16388
  PE6-mloc
  SvcId: 2 Name: Epipe-2
2001:bbbb:6:4005::                              End.b6.encaps* 16389
  PE6-mloc
  None
2001:bbbb:6:4010::                              uA             16400
  PE6-mloc
  None
-------------------------------------------------------------------------------
SIDs : 5
-------------------------------------------------------------------------------
===============================================================================
* indicates that the corresponding row element may have been truncated.
```

The following command shows the details of the SRv6 policy tunnel on head-end PE-6:

```
[/]
A:admin@PE-6# show router tunnel-table ipv6 protocol srv6-policy detail

===============================================================================
Tunnel Table (Router: Base)
===============================================================================
Destination      : 2001:db8::2:1/128
NextHop          : fpe_1.a
NextHop Weight   : 1
Tunnel Flags     : has-color
Age              : 00h01m40s            Color            : 200
CBF Classes      : (Not Specified)
Owner            : srv6-pol             Encap            : SRV6
Tunnel ID        : 917506               Preference       : 14
Tunnel SRV6 SID  : 2001:bbbb:5:4000:4:4 Tunnel Metric    : 0
                   000:3:4000
Tunnel MTU       :  -                   Max Label Stack  : 3
-------------------------------------------------------------------------------
Number of tunnel-table entries          : 1
Number of tunnel-table entries with LFA : 0
===============================================================================
```

The tunnel SRv6 SID in the preceding output shows that the micro-segment SIDs were compressed. The three micro-SIDs (2001:bbbb:5:4000::, 2001:bbbb:4:4000::, and 2001:bbbb:3:4000::) were compressed by the head-end to form a single container: 2001:bbbb:5:4000:4:4000:3:4000. This compression was, in part, possible because the three micro-SIDs belong to the same SID block.

The SRv6 policy tunnel to 2001:db8::2:1 with color 200 has tunnel ID 917506 on head-end PE-6. The following IPv6 routes use the SRv6 policy tunnel with ID 917506:

```
[/]
A:admin@PE-6# show router route-table ipv6 protocol srv6-policy

===============================================================================
IPv6 Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
```

```
-------------------------------------------------------------------------------
2001:bbbb:6:4005::/64                       Local   SRV6-Pol* 00h02m22s  14
       2001:db8::2:1 (tunneled:SRV6-Policy:917506)                  1
2001:bbbb:4005::/48                         Local   SRV6-Pol* 00h02m22s  14
       2001:db8::2:1 (tunneled:SRV6-Policy:917506)                  1
-------------------------------------------------------------------------------
No. of Routes: 2
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
===============================================================================
* indicates that the corresponding row element may have been truncated.
```

The following command shows that the BGP next hop 2001:db8::2:1 is resolved to an SRv6 policy with color 200:

```
[/]
A:admin@PE-6# show router bgp next-hop evpn 2001:db8::2:1 detail
===============================================================================
 BGP Router ID:192.0.2.6        AS:64500       Local AS:64500
===============================================================================

===============================================================================
BGP VPN Next Hop
===============================================================================
-------------------------------------------------------------------------------
---snip--- ## VPN next hop for Epipe-1 (no SRv6 policy; only GRE/RTM)

-------------------------------------------------------------------------------
VPN Next Hop          : 2001:db8::2:1
Autobind              : gre/rtm srv6-policy
Labels                : --
User-labels           : 1
Admin-tag-policy      : --
Strict-tunnel-tagging : N
Color                 : 200
UPA Trigger Next Hop  : --
Locator               : 2001:bbbb:1::/48
Created               : 00h03m50s
Last-modified         : 00h03m17s
-------------------------------------------------------------------------------
Resolving Prefix : 2001:db8::2:1/128
Preference       : 14                   Metric           : 0
Reference Count  : 1                    Owner            : SRV6-POLICY
Fib Programmed   : Y
Resolved Next Hop: 0.140.1.1
Egress Label     : n/a                  TunnelId         : 917506
Locator State    : Resolved
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
Next Hops : 2
===============================================================================
```

The following command on PE-6 shows the SRv6 destinations in Epipe-2, where 2001:bbbb:1:4004:: is the uDX2 SID on endpoint PE-1:

```
[/]
A:admin@PE-6# show service id 2 segment-routing-v6 destinations

===============================================================================
TEP, SID (Instance 1)
```

```
===============================================================================
TEP Address                             Segment Id
-------------------------------------------------------------------------------
2001:db8::2:1                           2001:bbbb:1:4004::
-------------------------------------------------------------------------------
Number of TEP, SID: 1
-------------------------------------------------------------------------------
===============================================================================

===============================================================================
Segment Routing v6 Ethernet Segment Dest (Instance 1)
===============================================================================
Eth SegId                                Num. Macs     Last Update
-------------------------------------------------------------------------------
No Matching Entries
===============================================================================
```

EVPN VPLS

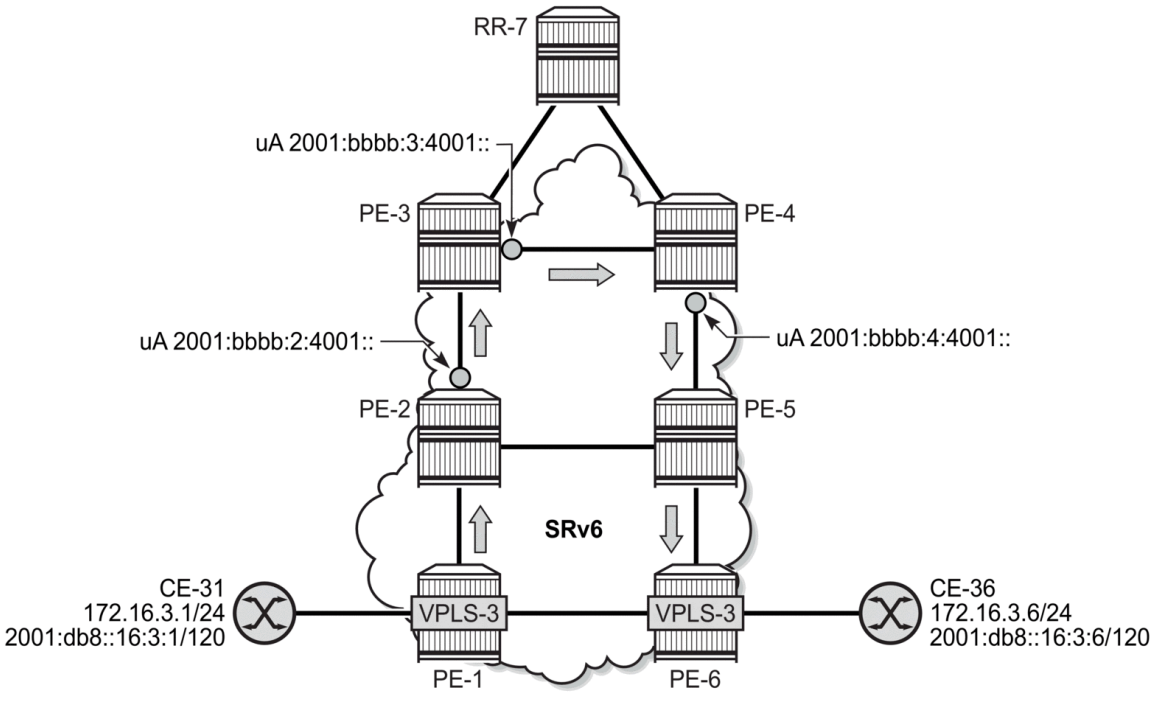
The following examples are described in this section:

- EVPN VPLS with static SRv6 policy on head-end PE-1
- EVPN VPLS with BGP-signaled SRv6 policy on head-end PE-6

EVPN VPLS with static SRv6 policy on head-end PE-1

Figure 73: EVPN VPLS using SRv6 policy with color 300 from PE-1 to PE-6 shows the uA SIDs in the segment list of a static SRv6 with color 300 on head-end PE-1.

Figure 73: EVPN VPLS using SRv6 policy with color 300 from PE-1 to PE-6



Static SRv6 policy

On head-end PE-1, a static SRv6 policy is configured with color 300 and endpoint 2001:db8::2:6, as follows:

```
# on head-end PE-1:
configure {
    router "Base" {
        segment-routing {
            sr-policies {
                static-policy "color-300-PE-1-PE-6-u" {
                    admin-state enable
                    color 300
                    endpoint 2001:db8::2:6
                    head-end local
                    type srv6
                    segment-routing-v6 {
                        binding-sid 1 {
                            micro-segment-locator {
                                locator-name "PE1-mloc"
                                function end-b6-encaps-red-next-csid
                            }
                        }
                    }
                    segment-list 1 {
                        admin-state enable
                        segment 1 {
                            srv6-sid 2001:bbbb:2:4001::  # int-PE-2-PE-3
```

```
                        behavior-and-structure {
                            behavior end-x-next-csid
                            block-length 32
                            node-length 16
                            function-length 16
                        }
                    }
                    segment 2 {
                        srv6-sid 2001:bbbb:3:4001::  # int-PE-3-PE-4
                        behavior-and-structure {
                            behavior end-x-next-csid
                            block-length 32
                            node-length 16
                            function-length 16
                        }
                    }
                    segment 3 {
                        srv6-sid 2001:bbbb:4:4001::  # int-PE-4-PE-5
                        behavior-and-structure {
                            behavior end-x-next-csid
                            block-length 32
                            node-length 16
                            function-length 16
                        }
                    }
                }
            }
```

EVPN VPLS

On PE-1, EVPN VPLS "VPLS-3" is configured as follows:

```
# on PE-1:
configure {
    service {
        vpls "VPLS-3" {
            admin-state enable
            service-id 3
            customer "1"
            segment-routing-v6 1 {
                micro-segment-locator "PE1-mloc" {
                    function {
                        udt2m {
                        }
                        udt2u {
                        }
                    }
                }
            }
            bgp 1 {
                route-target {
                    export "target:64500:3"
                    import "target:64500:3"
                }
            }
            bgp-evpn {
                evi 3
                segment-routing-v6 1 {
                    admin-state enable
                    source-address 2001:db8::2:1
                    resolution fallback-tunnel-to-route-table
```

```
                    srv6 {
                        instance 1
                        default-locator "PE1-mloc"
                    }
                    route-next-hop {
                        ip-address 2001:db8::2:1
                    }
                }
            }
            sap 1/1/c10/1:3 {
                description "SAP to CE-31"
            }
        }
```

PE-6 exports color 300 which matches the color in the static SRv6 policy "color-300-PE-1-PE-6-u" on PE-1, as follows:

```
#  on PE-6:
configure exclusive
    policy-options {
        community "color-300" {
            member "color:00:300" { }
        }
        community "vsi-3" {
            member "target:64500:3" { }
        }
        policy-statement "vpls-3-export-c300" {
            default-action {
                action-type accept
                community {
                    add ["vsi-3" "color-300"]
                }
            }
        }
    }
    service {
        vpls "VPLS-3" {
            admin-state enable
            service-id 3
            customer "1"
            segment-routing-v6 1 {
                micro-segment-locator "PE6-mloc" {
                    function {
                        udt2m {
                        }
                        udt2u {
                        }
                    }
                }
            }
            bgp 1 {
                vsi-export ["vpls-3-export-c300"]
                route-target {
                    import "target:64500:3"
                }
            }
            bgp-evpn {
                evi 3
                segment-routing-v6 1 {
                    admin-state enable
                    source-address 2001:db8::2:6
                    srv6 {
                        instance 1
```

```
                    default-locator "PE6-mloc"
                }
                route-next-hop {
                    ip-address 2001:db8::2:6
                }
            }
        }
        sap 1/1/c10/1:3 {
            description "SAP to CE-36"
        }
    }
```

Verification

PE-1 receives the following IMET route from PE-6. The color:00:300 community matches the color 300 in the static SRv6 policy.

```
[/]
A:admin@PE-1# show router bgp routes evpn incl-mcast rd 192.0.2.6:3 detail
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================
 Legend -
 Status codes  : u - used, s - suppressed, h - history, d - decayed, * - valid
                 l - leaked, x - stale, > - best, b - backup, p - purge
 Origin codes  : i - IGP, e - EGP, ? - incomplete

===============================================================================
BGP EVPN Inclusive-Mcast Routes
===============================================================================
Original Attributes

Network        : n/a
Nexthop        : 2001:db8::2:6
Path Id        : None
From           : 2001:db8::2:7
Res. Nexthop   : fe80::24:1ff:fe01:1
Local Pref.    : 100                    Interface Name : int-PE-1-PE-6
Aggregator AS  : None                   Aggregator     : None
Atomic Aggr.   : Not Atomic             MED            : None
AIGP Metric    : None                   IGP Cost       : 10
Connector      : None
Community      : target:64500:3 color:00:300
Cluster        : 192.0.2.7
Originator Id  : 192.0.2.6              Peer Router Id : 192.0.2.7
Origin         : IGP
Flags          : Used Valid Best
Route Source   : Internal
AS-Path        : No As-Path
EVPN type      : INCL-MCAST
Tag            : 0
Originator IP  : 2001:db8::2:6
Route Dist.    : 192.0.2.6:3
Route Tag      : 0
Neighbor-AS    : n/a
DB Orig Val    : N/A                    Final Orig Val : N/A
Source Class   : 0                      Dest Class     : 0
Add Paths Send : Default
Last Modified  : 00h00m27s
SRv6 TLV Type  : SRv6 L2 Service TLV (6)
SRv6 SubTLV    : SRv6 SID Information (1)
```

```
Sid            : 2001:bbbb:6::
Full Sid       : 2001:bbbb:6:4012::
Behavior       : End.uDT2M (68)
SRv6 SubSubTLV : SRv6 SID Structure (1)
Loc-Block-Len  : 32                    Loc-Node-Len   : 16
Func-Len       : 16                    Arg-Len        : 16
Tpose-Len      : 16                    Tpose-offset   : 48
-------------------------------------------------------------------------------
PMSI Tunnel Attributes :
Tunnel-type    : Ingress Replication
Flags          : Type: RNVE(0) BM: 0 U: 0 Leaf: not required
MPLS Label     : 4198912
Tunnel-Endpoint: 2001:db8::2:6
-------------------------------------------------------------------------------
---snip---
```

When traffic has been sent between CE-31 and CE-36, the forwarding database (FDB) on PE-1 is as follows:

```
[/]
A:admin@PE-1# show service id "VPLS-3" fdb detail

===============================================================================
Forwarding Database, Service 3
===============================================================================
ServId     MAC               Source-Identifier       Type     Last Change
           Transport:Tnl-Id                          Age
-------------------------------------------------------------------------------
3          00:00:5e:00:53:31 sap:1/1/c10/1:3         L/0      06/13/24 13:42:27
3          00:00:5e:00:53:36 srv6-1:                 Evpn     06/13/24 13:42:34
                             2001:db8::2:6
           2001:bbbb:6:4011::
-------------------------------------------------------------------------------
No. of MAC Entries: 2
-------------------------------------------------------------------------------
Legend:L=Learned O=Oam P=Protected-MAC C=Conditional S=Static Lf=Leaf T=Trusted
===============================================================================
```

MAC address 00:00:5e:00:53:36 from CE-36 can be reached through an SRv6 policy tunnel. The following IPv6 tunnel table shows that the SRv6 policy tunnel to 2001:db8::2:6/128 with color 300 has ID 917508:

```
[/]
A:admin@PE-1# show router tunnel-table ipv6 protocol srv6-policy

===============================================================================
IPv6 Tunnel Table (Router: Base)
===============================================================================
Destination                                 Owner     Encap TunnelId  Pref
Nexthop                                     Color           Metric
-------------------------------------------------------------------------------
2001:db8::2:6/128                           srv6-pol  SRV6  917506    14
  fpe_1.a                                     100             0
2001:db8::2:6/128                           srv6-pol  SRV6  917508    14
  fpe_1.a                                     300             0
-------------------------------------------------------------------------------
Flags: B = BGP or MPLS backup hop available
       L = Loop-Free Alternate (LFA) hop available
       E = Inactive best-external BGP route
       k = RIB-API or Forwarding Policy backup hop
===============================================================================
```

The following command shows the micro-segment local SIDs. Besides the node SID uN and the adjacency SID uA, SIDs for different services are listed. SID 2001:bbbb:1:4004:: with uDX2 function applies to EVPN VPWS Epipe-2; SID 2001:bbbb:1:4011:: with uDT2U function and SID 2001:bbbb:1:4012:: with uDT2M function apply to EVPN VPLS-3. The End.B6.Encaps.Red SID 2001:bbbb:1:43e8:: is the binding SID for the SRv6 policy "color-300-PE-1-PE-6-u" which is used in VPLS-3.

```
[/]
A:admin@PE-1# show router segment-routing-v6 micro-segment-local-sid

===============================================================================
Micro Segment Routing v6 Local SIDs
===============================================================================
SID                                                Type           Function
  Micro Segment Locator
  Context
-------------------------------------------------------------------------------
2001:bbbb:1::                                      uN             1
  PE1-mloc
  None
2001:bbbb:1:4000::                                 uA             16384
  PE1-mloc
  Base
2001:bbbb:1:4004::                                 uDX2           16388
  PE1-mloc
  SvcId: 2 Name: Epipe-2
2001:bbbb:1:4010::                                 uA             16400
  PE1-mloc
  None
2001:bbbb:1:4011::                                 uDT2U          16401
  PE1-mloc
  SvcId: 3 Name: VPLS-3
2001:bbbb:1:4012::                                 uDT2M          16402
  PE1-mloc
  SvcId: 3 Name: VPLS-3
2001:bbbb:1:43e8::                                 End.b6.encaps* 17384
  PE1-mloc
  None
-------------------------------------------------------------------------------
SIDs : 7
-------------------------------------------------------------------------------
===============================================================================
* indicates that the corresponding row element may have been truncated.
```

The IPv6 route table shows two routes that use the SRv6 policy tunnel with tunnel ID 917508:

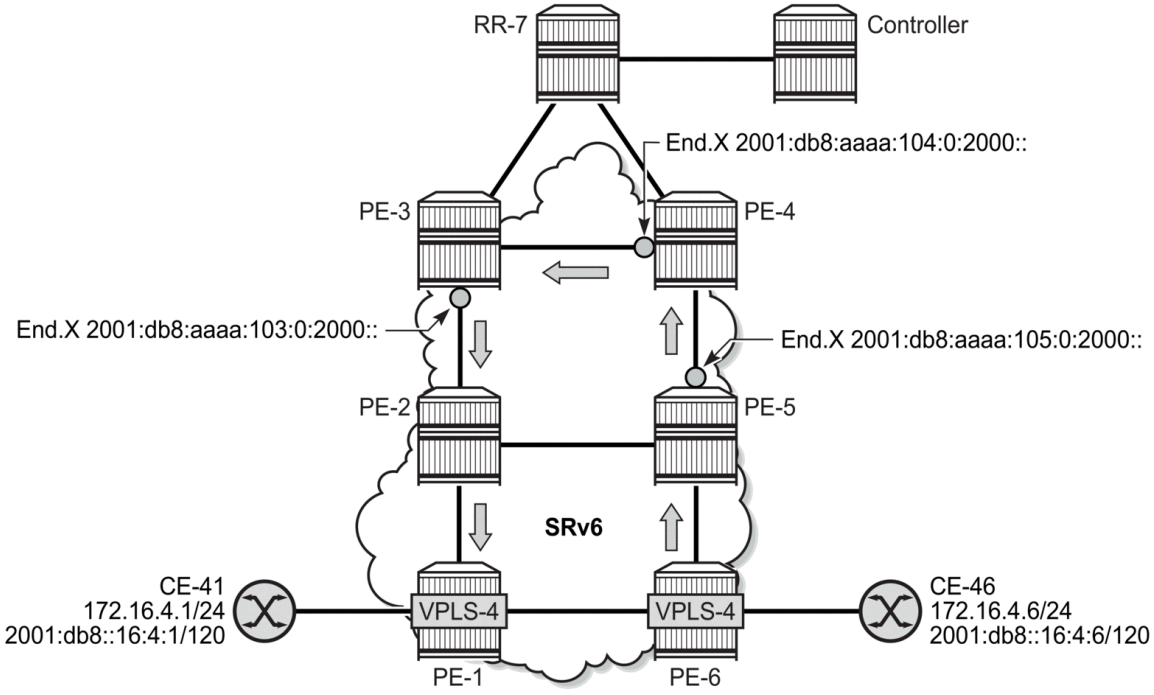
```
[/]
A:admin@PE-1# show router route-table ipv6 protocol srv6-policy

===============================================================================
IPv6 Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
2001:db8:aaaa:101:1:2000::/128                Local   SRV6-Pol* 00h37m48s  14
       2001:db8::2:6 (tunneled:SRV6-Policy:917506)                  1
2001:bbbb:1:43e8::/64                         Local   SRV6-Pol* 00h03m45s  14
       2001:db8::2:6 (tunneled:SRV6-Policy:917508)                  1
2001:bbbb:43e8::/48                           Local   SRV6-Pol* 00h03m45s  14
       2001:db8::2:6 (tunneled:SRV6-Policy:917508)                  1
-------------------------------------------------------------------------------
No. of Routes: 3
```

```
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
===============================================================================
* indicates that the corresponding row element may have been truncated.
```

EVPN VPLS with BGP-signaled SRv6 policy on head-end PE-6

Figure 74: EVPN VPLS using SRv6 policy with color 400 from PE-6 to PE-1 shows the End.X SIDs in the segment list of a BGP-signaled SRv6 policy on head-end PE-6. This SRv6 policy is configured on the controller, imported in the BGP RIB, and advertised as an SRv6 policy route in BGP.

Figure 74: EVPN VPLS using SRv6 policy with color 400 from PE-6 to PE-1



BGP-signaled SRv6 policy

The following SRv6 policy with head-end 192.0.2.6 is configured on the controller and is advertised as an SRv6 policy:

```
# on controller:
configure exclusive
    router "Base" {
        segment-routing {
            sr-policies {
                admin-state enable
                static-policy "color-400-PE-6-PE-1" {
```

```
                    admin-state enable
                    color 400
                    endpoint 2001:db8::2:1
                    head-end 192.0.2.6
                    distinguisher 400006001
                    type srv6
                    segment-routing-v6 {
                        binding-sid 1 {
                            ip-address 2001:db8:aaaa:106:0:9000::
                        }
                    }
                    segment-list 1 {
                        admin-state enable
                        segment 1 {
                            srv6-sid 2001:db8:aaaa:105:0:2000::   # "int-PE-5-PE-4"
                        }
                        segment 2 {
                            srv6-sid 2001:db8:aaaa:104:0:2000::   # "int-PE-4-PE-3"
                        }
                        segment 3 {
                            srv6-sid 2001:db8:aaaa:103:0:2000::   # "int-PE-3-PE-2"
                        }
                    }
                }
```

EVPN VPLS on PE-1 and PE-6

The EVPN VPLS "VPLS-4" is configured as follows. PE-1 exports color 400 which matches the color in the policy "color-400-PE-6-PE-1" on head-end PE-6:

```
# on PE-1:
configure {
    policy-options {
        community "color-400" {
            member "color:00:400" { }
        }
        community "vsi-4" {
            member "target:64500:4" { }
        }
        policy-statement "vpls-4-export-c400" {
            default-action {
                action-type accept
                community {
                    add ["vsi-4" "color-400"]
                }
            }
        }
    }
    service {
        vpls "VPLS-4" {
            admin-state enable
            service-id 4
            customer "1"
            segment-routing-v6 1 {
                locator "PE1-loc" {
                    function {
                        end-dt2u {
                        }
                        end-dt2m {
                        }
                    }
```

```
            }
        }
        bgp 1 {
            vsi-export ["vpls-4-export-c400"]
            route-target {
                import "target:64500:4"
            }
        }
        bgp-evpn {
            evi 4
            segment-routing-v6 1 {
                admin-state enable
                source-address 2001:db8::2:1
                srv6 {
                    instance 1
                    default-locator "PE1-loc"
                }
                route-next-hop {
                    ip-address 2001:db8::2:1
                }
            }
        }
        sap 1/1/c10/1:4 {
            description "SAP to CE-41"
        }
    }
```

On head-end PE-6, only the route target is exported, not the color. The EVPN VPLS "VPLS-4" is configured with SRv6 resolution fallback from the tunnel table to the route table, as follows:

```
#  on PE-6:
configure {
    service {
        vpls "VPLS-4" {
            admin-state enable
            service-id 4
            customer "1"
            segment-routing-v6 1 {
                locator "PE6-loc" {
                    function {
                        end-dt2u {
                        }
                        end-dt2m {
                        }
                    }
                }
            }
            bgp 1 {
                route-target {
                    export "target:64500:4"
                    import "target:64500:4"
                }
            }
            bgp-evpn {
                evi 4
                segment-routing-v6 1 {
                    admin-state enable
                    source-address 2001:db8::2:6
                    resolution fallback-tunnel-to-route-table
                    srv6 {
                        instance 1
                        default-locator "PE6-loc"
                    }
```

```
                    route-next-hop {
                        ip-address 2001:db8::2:6
                    }
                }
            }
            sap 1/1/c10/1:4 {
                description "SAP to CE-46"
            }
        }
```

Verification

On head-end PE-6, the BGP-signaled SRv6 with color 400 is as follows:

```
[/]
A:admin@PE-6# show router segment-routing sr-policies bgp color 400

===============================================================================
SR-Policies Path
===============================================================================
-------------------------------------------------------------------------------
Type            : srv6
Active          : Yes                   Owner           : bgp
Operational     : Yes
Color           : 400
Head            : 0.0.0.0               Endpoint Addr   : 2001:db8::2:1
RD              : 400006001             Preference      : 100
SRv6 BSID 1     : 2001:db8:aaaa:106:0:9000::
TunnelId        : 917507                Age             : 42
Origin ASN      : 64500                 Origin          : 2001:db8::2:8
NumReEval       : 0                     LastReEvalReason: none
NumActPathChange: 0                     Last Change     : 06/13/2024 13:47:57
Maintenance Plcy:
Ret Path BFD SID:

Path Segment Lists:
Segment-List    : 1                     Weight          : 1
Num Segments    : 3                     Last Change     : 06/13/2024 12:58:00
  1 SRv6 SID  : 2001:db8:aaaa:105:0:2000::               State : resolved-up
  2 SRv6 SID  : 2001:db8:aaaa:104:0:2000::               State : N/A
  3 SRv6 SID  : 2001:db8:aaaa:103:0:2000::               State : N/A

===============================================================================
```

The second SRv6 policy tunnel in the IPv6 tunnel table has color 400:

```
[/]
A:admin@PE-6# show router tunnel-table ipv6 protocol srv6-policy

===============================================================================
IPv6 Tunnel Table (Router: Base)
===============================================================================
Destination                                 Owner     Encap TunnelId  Pref
Nexthop                                     Color           Metric
-------------------------------------------------------------------------------
2001:db8::2:1/128                           srv6-pol  SRV6  917506    14
  fpe_1.a                                     200             0
2001:db8::2:1/128                           srv6-pol  SRV6  917507    14
  fpe_1.a                                     400             0
-------------------------------------------------------------------------------
```

```
Flags: B = BGP or MPLS backup hop available
       L = Loop-Free Alternate (LFA) hop available
       E = Inactive best-external BGP route
       k = RIB-API or Forwarding Policy backup hop
===============================================================================
```

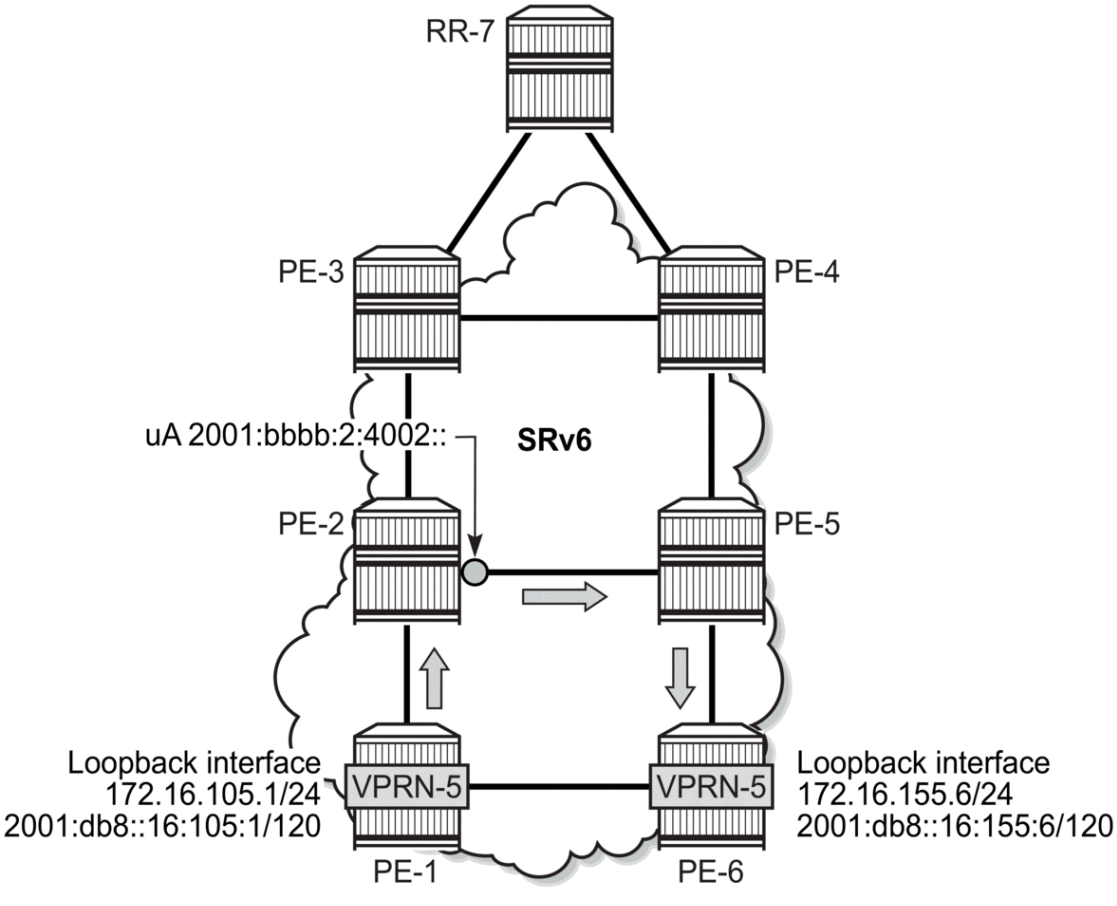
The BGP next hop for EVPN routes to endpoint 2001:db8::2:1 is resolved using the SRv6 policy tunnel with color 400:

```
[/]
A:admin@PE-6# show router bgp next-hop evpn 2001:db8::2:1 detail
===============================================================================
 BGP Router ID:192.0.2.6        AS:64500       Local AS:64500
===============================================================================

===============================================================================
BGP VPN Next Hop
===============================================================================
-------------------------------------------------------------------------------
---snip---

-------------------------------------------------------------------------------
VPN Next Hop          : 2001:db8::2:1
Autobind              : gre/rtm srv6-policy
Labels                : --
User-labels           : 1
Admin-tag-policy      : --
Strict-tunnel-tagging : N
Color                 : 400
UPA Trigger Next Hop  : --
Locator               : 2001:db8:aaaa:101::/64
Created               : 00h03m48s
Last-modified         : 00h02m54s
-------------------------------------------------------------------------------
Resolving Prefix : 2001:db8::2:1/128
Preference       : 14                   Metric           : 0
Reference Count  : 2                    Owner            : SRV6-POLICY
Fib Programmed   : Y
Resolved Next Hop: 0.140.1.1
Egress Label     : n/a                  TunnelId         : 917507
Locator State    : Resolved
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
Next Hops : 4
===============================================================================
```

EVPN IFL

Figure 75: EVPN IFL using SRv6 policy with color 500 from PE-1 to PE-6 shows the micro-segment adjacency SID (uA) for the interface "int-PE-2-PE-5" on PE-2, which is the only segment in the segment list of a static SRv6 policy that is configured on head-end PE-1.

Figure 75: EVPN IFL using SRv6 policy with color 500 from PE-1 to PE-6



Static SRv6 policy on PE-1

On head-end PE-1, the SRv6 policy with color 500 is configured as follows:

```
# on PE-1:
configure {
    router "Base" {
        segment-routing {
            sr-policies {
                admin-state enable
                static-policy "color-500-PE-1-PE-6-u" {
                    admin-state enable
                    color 500
                    endpoint 2001:db8::2:6
                    head-end local
                    type srv6
                    segment-routing-v6 {
                        binding-sid 1 {
                            micro-segment-locator {
                                locator-name "PE1-mloc"
```

```
                            function end-b6-encaps-red-next-csid
                        }
                    }
                }
                segment-list 1 {
                    admin-state enable
                    segment 1 {
                        srv6-sid 2001:bbbb:2:4002::     # uA "int-PE-2-PE-5"
                        behavior-and-structure {
                            behavior end-x-next-csid
                            block-length 32
                            node-length 16
                            function-length 16
                        }
                    }
                }
            }
```

VPRN-5 with EVPN IFL

VPRN-5 is configured with BGP EVPN, as follows:

```
# on PE-1:
configure {
    service {
        vprn "VPRN-5" {
            admin-state enable
            service-id 5
            customer "1"
            autonomous-system 64496
            segment-routing-v6 1 {
                micro-segment-locator "PE1-mloc" {
                    function {
                        udt46 {
                        }
                    }
                }
            }
            bgp-evpn {
                segment-routing-v6 1 {
                    admin-state enable
                    route-distinguisher "192.0.2.1:5"
                    source-address 2001:db8::2:1
                    resolution fallback-tunnel-to-route-table
                    vrf-target {
                        community "target:64500:5"
                    }
                    srv6 {
                        instance 1
                        default-locator "PE1-mloc"
                    }
                }
            }
            interface "lo1" {
                loopback true
                ipv4 {
                    primary {
                        address 172.16.105.1
                        prefix-length 24
                    }
                }
```

```
                ipv6 {
                    address 2001:db8::105:1 {
                        prefix-length 120
                    }
                }
            }
```

Endpoint PE-6 exports color 500. The configuration on PE-6 is as follows:

```
# on PE-6:
configure {
    policy-options {
        community "color-500" {
            member "color:00:500" { }
        }
        community "vrf-5" {
            member "target:64500:5" { }
        }
        policy-statement "vrf-5-export-c500" {
            default-action {
                action-type accept
                community {
                    add ["vrf-5" "color-500"]
                }
            }
        }
        policy-statement "vrf-5-import" {
            entry 10 {
                from {
                    community {
                        name "vrf-5"
                    }
                }
                action {
                    action-type accept
                }
            }
        }
    }
    service {
        vprn "VPRN-5" {
            admin-state enable
            service-id 5
            customer "1"
            autonomous-system 64497
            segment-routing-v6 1 {
                micro-segment-locator "PE6-mloc" {
                    function {
                        udt46 {
                        }
                    }
                }
            }
            bgp-evpn {
                segment-routing-v6 1 {
                    admin-state enable
                    route-distinguisher "192.0.2.1:5"
                    source-address 2001:db8::2:6
                    vrf-import {
                        policy ["vrf-5-import"]
                    }
                    vrf-export {
                        policy ["vrf-5-export-c500"]
```

```
                    }
                    srv6 {
                        instance 1
                        default-locator "PE6-mloc"
                    }
                }
            }
            interface "lo1" {
                loopback true
                ipv4 {
                    primary {
                        address 172.16.155.6
                        prefix-length 24
                    }
                }
                ipv6 {
                    address 2001:db8::155:6 {
                        prefix-length 120
                    }
                }
            }
        }
```

The BGP configuration in the Base router is modified to ensure that IPv6 next hops are advertised for EVPN routes, as described in the BGP configuration section.

Verification

PE-1 receives the EVPN IP prefix 172.16.155.0/24 and the EVPN IPv6 prefix 2001:db8::155:0/120 from PE-6, as follows:

```
[/]
A:admin@PE-1# show router bgp routes evpn ip-prefix
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================
 Legend -
 Status codes  : u - used, s - suppressed, h - history, d - decayed, * - valid
                 l - leaked, x - stale, > - best, b - backup, p - purge
 Origin codes  : i - IGP, e - EGP, ? - incomplete

===============================================================================
BGP EVPN IP-Prefix Routes
===============================================================================
Flag  Route Dist.         Prefix
      Tag                 Gw Address
                          NextHop
                          Label
                          ESI
-------------------------------------------------------------------------------
u*>i  192.0.2.1:5         172.16.155.0/24
      0                   00:00:00:00:00:00
                          2001:db8::2:6
                          16403
                          ESI-0

-------------------------------------------------------------------------------
Routes : 1
===============================================================================
```

```
[/]
A:admin@PE-1# show router bgp routes evpn ipv6-prefix
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================
 Legend -
 Status codes  : u - used, s - suppressed, h - history, d - decayed, * - valid
                 l - leaked, x - stale, > - best, b - backup, p - purge
 Origin codes  : i - IGP, e - EGP, ? - incomplete

===============================================================================
BGP EVPN IPv6-Prefix Routes
===============================================================================
Flag  Route Dist.         Prefix
      Tag                 Gw Address
                          NextHop
                          Label
                          ESI
-------------------------------------------------------------------------------
u*>i  192.0.2.1:5         2001:db8::155:0/120
      0                   00:00:00:00:00:00
                          2001:db8::2:6
                          16403
                          ESI-0

-------------------------------------------------------------------------------
Routes : 1
===============================================================================
```

In this example, multiple services use different SRv6 policies and the list of micro-segment local SIDs for VPRN-5 shows multiple SRv6 SIDs for the services. On head-end PE-1, the list of micro-segment local SIDs shows the uDT46 function and the End.B6.Encaps.Red function for two different SRv6 policy tunnels: SID 2001:bbbb:1:43e9:: corresponds to the SRv6 policy tunnel for color 500; SID 2001:bbbb:1:43e8:: corresponds to the SRv6 policy tunnel for color 300 and is not used for VPRN-5.

```
[/]
A:admin@PE-1# show router segment-routing-v6 micro-segment-local-sid context "VPRN-5"

===============================================================================
Micro Segment Routing v6 Local SIDs
===============================================================================
SID                                             Type           Function
  Micro Segment Locator
  Context
-------------------------------------------------------------------------------
2001:bbbb:1:4013::                              uDT46          16403
  PE1-mloc
  SvcId: 5 Name: VPRN-5
2001:bbbb:1:43e8::                              End.b6.encaps* 17384
  PE1-mloc
  None
2001:bbbb:1:43e9::                              End.b6.encaps* 17385
  PE1-mloc
  None
-------------------------------------------------------------------------------
SIDs : 3
-------------------------------------------------------------------------------
===============================================================================
* indicates that the corresponding row element may have been truncated.
```

If services use different SRv6 policies, the next hop command for the service shows multiple SRv6 policies for the services, but the traffic flow uses the right SRv6 policy with the corresponding color.

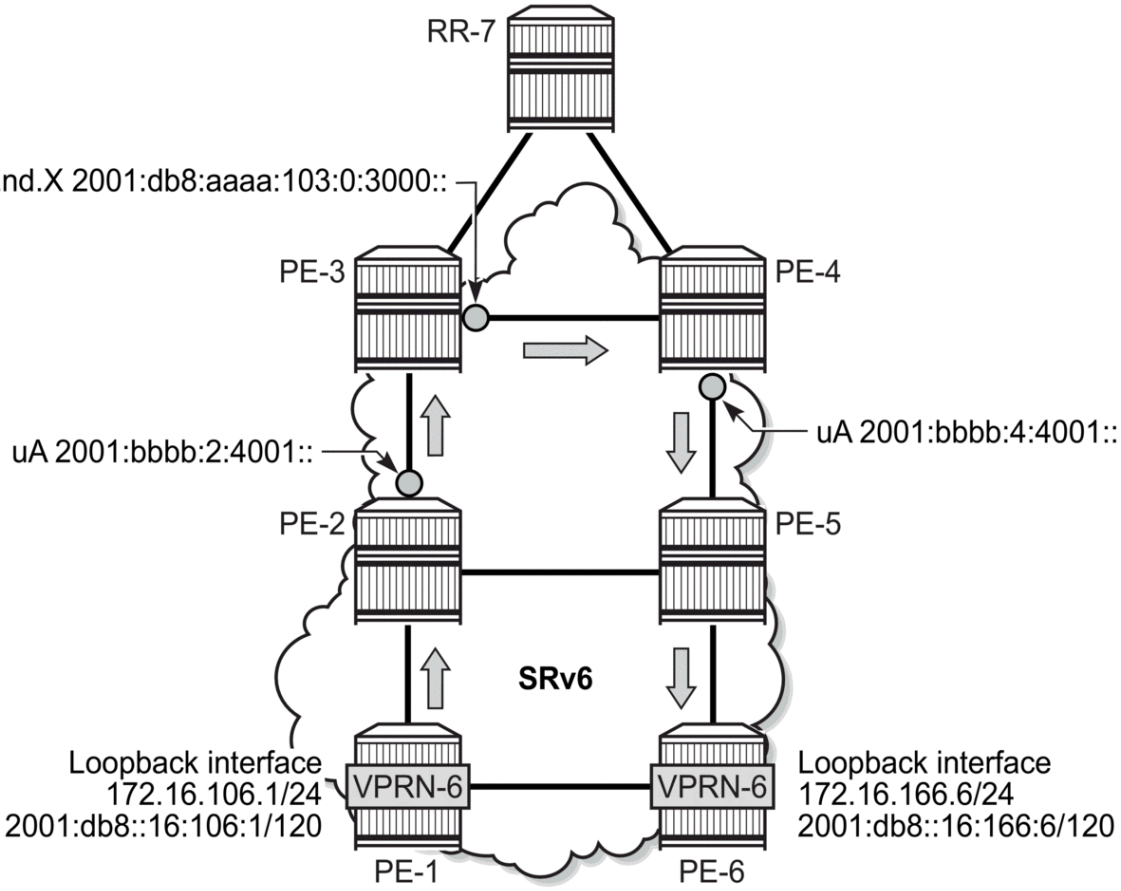
```
[/]
A:admin@PE-1# show router bgp next-hop 2001:db8::2:6 evpn detail
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================

===============================================================================
BGP VPN Next Hop
===============================================================================
-------------------------------------------------------------------------------
---snip---      # GRE tunnels for Epipe-2 and VPLS-4 are not shown

-------------------------------------------------------------------------------
VPN Next Hop          : 2001:db8::2:6
Autobind              : gre/rtm srv6-policy
Labels                : --
User-labels           : 1
Admin-tag-policy      : --
Strict-tunnel-tagging : N
Color                 : 100
UPA Trigger Next Hop  : --
Locator               : 2001:db8:aaaa:106::/64
Created               : 00h02m45s
Last-modified         : 00h02m45s
-------------------------------------------------------------------------------
Resolving Prefix : 2001:db8::2:6/128
Preference       : 14                   Metric           : 0
Reference Count  : 1                    Owner            : SRV6-POLICY
Fib Programmed   : Y
Resolved Next Hop: 0.140.1.1
Egress Label     : n/a                  TunnelId         : 917506
Locator State    : Resolved
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
VPN Next Hop          : 2001:db8::2:6
Autobind              : gre/rtm srv6-policy
Labels                : --
User-labels           : 1
Admin-tag-policy      : --
Strict-tunnel-tagging : N
Color                 : 300
UPA Trigger Next Hop  : --
Locator               : 2001:bbbb:6::/48
Created               : 00h02m45s
Last-modified         : 00h02m45s
-------------------------------------------------------------------------------
Resolving Prefix : 2001:db8::2:6/128
Preference       : 14                   Metric           : 0
Reference Count  : 1                    Owner            : SRV6-POLICY
Fib Programmed   : Y
Resolved Next Hop: 0.140.1.1
Egress Label     : n/a                  TunnelId         : 917508
Locator State    : Resolved
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
VPN Next Hop          : 2001:db8::2:6
Autobind              : gre srv6-policy
Labels                : --
User-labels           : 2
Admin-tag-policy      : --
```

```
Strict-tunnel-tagging : N
Color                 : 500
UPA Trigger Next Hop  : --
Locator               : 2001:bbbb:6::/48
Created               : 00h02m45s
Last-modified         : 00h02m45s
-------------------------------------------------------------------------------
Resolving Prefix : 2001:db8::2:6/128
Preference       : 14                   Metric           : 0
Reference Count  : 2                    Owner            : SRV6-POLICY
Fib Programmed   : Y
Resolved Next Hop: 0.140.1.1
Egress Label     : n/a                  TunnelId         : 917509
Locator State    : Resolved
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
Next Hops : 5
===============================================================================
```

IP VPN

Figure 76: IP VPN using SRv6 policy with color 600 from PE-1 to PE-6 shows a combination of classic adjacency SIDs (End.X) and micro-segment adjacency SIDs (uA) in the segment list of a static SRv6 policy that is configured on head-end PE-1.

Figure 76: IP VPN using SRv6 policy with color 600 from PE-1 to PE-6



The BGP configuration for IP VPN is described in the BGP configuration section.

Static SRv6 policy

On head-end 1, the following SRv6 policy is configured with color 600 and a segment list containing two micro-segment SRv6 SIDs and one classic SRv6 SID.

```
# on PE-1:
configure {
    router "Base" {
        segment-routing {
            sr-policies {
                admin-state enable
                static-policy "color-600-PE-1-PE-6-mixed" {
                    admin-state enable
                    color 600
                    endpoint 2001:db8::2:6
                    head-end local
                    type srv6
```

```
                    segment-routing-v6 {
                        binding-sid 1 {
                            micro-segment-locator {
                                locator-name "PE1-mloc"
                                function end-b6-encaps-red-next-csid
                            }
                        }
                    }
                    segment-list 1 {
                        admin-state enable
                        segment 1 {
                            srv6-sid 2001:bbbb:2:4001::
                            behavior-and-structure {
                                behavior end-x-next-csid
                                block-length 32
                                node-length 16
                                function-length 16
                            }
                        }
                        segment 2 {
                            srv6-sid 2001:db8:aaaa:103:0:3000::
                        }
                        segment 3 {
                            srv6-sid 2001:bbbb:4:4001::
                            behavior-and-structure {
                                behavior end-x-next-csid
                                block-length 32
                                node-length 16
                                function-length 16
                            }
                       }
                    }
                }
```

VPRN-6 with VPN IP

On head-end PE-1, VPRN-6 is configured as follows:

```
# on PE-1:
configure {
    service {
        vprn "VPRN-6" {
            admin-state enable
            service-id 6
            customer "1"
            autonomous-system 64496
            segment-routing-v6 1 {
                micro-segment-locator "PE1-mloc" {
                    function {
                        udt4 {
                        }
                        udt6 {
                        }
                    }
                }
            }
            bgp-ipvpn {
                segment-routing-v6 1 {
                    admin-state enable
                    route-distinguisher "192.0.2.1:6"
                    source-address 2001:db8::2:1
```

```
                    resolution fallback-tunnel-to-route-table
                    vrf-target {
                        community "target:64500:6"
                    }
                    srv6 {
                        instance 1
                        default-locator "PE1-mloc"
                    }
                }
            }
            interface "lo1" {
                loopback true
                ipv4 {
                    primary {
                        address 172.16.106.1
                        prefix-length 24
                    }
                }
                ipv6 {
                    address 2001:db8::106:1 {
                        prefix-length 120
                    }
                }
            }
```

On endpoint PE-6, the configuration is as follows:

```
# on PE-6:
configure {
    policy-options {
        community "color-600" {
            member "color:00:600" { }
        }
        community "vrf-6" {
            member "target:64500:6" { }
        }
        policy-statement "vrf-6-export-c600" {
            default-action {
                action-type accept
                community {
                    add ["vrf-6" "color-600"]
                }
            }
        }
        policy-statement "vrf-6-import" {
            entry 10 {
                from {
                    community {
                        name "vrf-6"
                    }
                }
                action {
                    action-type accept
                }
            }
        }
    }
    service {
        vprn "VPRN-6" {
            admin-state enable
            service-id 6
            customer "1"
            autonomous-system 64497
```

```
            segment-routing-v6 1 {
                micro-segment-locator "PE6-mloc" {
                    function {
                        udt4 {
                        }
                        udt6 {
                        }
                    }
                }
            }
            bgp-ipvpn {
                segment-routing-v6 1 {
                    admin-state enable
                    route-distinguisher "192.0.2.1:6"
                    source-address 2001:db8::2:6
                    vrf-import {
                        policy ["vrf-6-import"]
                    }
                    vrf-export {
                        policy ["vrf-6-export-c600"]
                    }
                    srv6 {
                        instance 1
                        default-locator "PE6-mloc"
                    }
                }
            }
            interface "lo1" {
                loopback true
                ipv4 {
                    primary {
                        address 172.16.166.6
                        prefix-length 24
                    }
                }
                ipv6 {
                    address 2001:db8::166:6 {
                        prefix-length 120
                    }
                }
            }
```

Verification

On head-end PE-1, the list of micro-segment local SIDs includes SID 2001:bbbb:1:43e8:: for VPLS-3 and SID 2001:bbbb:1:43e9:: for VPRN-5, but only SID 2001:bbbb:1:43ea:: is used for VPRN-6.

```
[/]
A:admin@PE-1# show router segment-routing-v6 micro-segment-local-sid context "VPRN-6"

===============================================================================
Micro Segment Routing v6 Local SIDs
===============================================================================
SID                                                 Type           Function
  Micro Segment Locator
  Context
-------------------------------------------------------------------------------
2001:bbbb:1:4014::                                  uDT4           16404
  PE1-mloc
  SvcId: 6 Name: VPRN-6
2001:bbbb:1:4015::                                  uDT6           16405
```

```
  PE1-mloc
  SvcId: 6 Name: VPRN-6
2001:bbbb:1:43e8::                                End.b6.encaps* 17384
  PE1-mloc
  None
2001:bbbb:1:43e9::                                End.b6.encaps* 17385
  PE1-mloc
  None
2001:bbbb:1:43ea::                                End.b6.encaps* 17386
  PE1-mloc
  None
-------------------------------------------------------------------------------
SIDs : 5
-------------------------------------------------------------------------------
===============================================================================
* indicates that the corresponding row element may have been truncated.
```

The tunnel with color 600 is the SRv6 policy tunnel to be used for VPRN-6:

```
[/]
A:admin@PE-1# show router tunnel-table ipv6 protocol srv6-policy

===============================================================================
IPv6 Tunnel Table (Router: Base)
===============================================================================
Destination                                     Owner     Encap TunnelId  Pref
Nexthop                                         Color           Metric
-------------------------------------------------------------------------------
2001:db8::2:6/128                               srv6-pol  SRV6  917506    14
  fpe_1.a                                         100             0
2001:db8::2:6/128                               srv6-pol  SRV6  917508    14
  fpe_1.a                                         300             0
2001:db8::2:6/128                               srv6-pol  SRV6  917509    14
  fpe_1.a                                         500             0
2001:db8::2:6/128                               srv6-pol  SRV6  917511    14
  fpe_1.a                                         600             0
-------------------------------------------------------------------------------
Flags: B = BGP or MPLS backup hop available
       L = Loop-Free Alternate (LFA) hop available
       E = Inactive best-external BGP route
       k = RIB-API or Forwarding Policy backup hop
===============================================================================
```

The BGP next hop for VPN-IPv4 routes is resolved using an SRv6 policy:

```
[/]
A:admin@PE-1# show router bgp next-hop vpn-ipv4
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================

===============================================================================
BGP VPN Next Hop
===============================================================================
VPN Next Hop                                                 Owner
   Autobind                                    FibProg       Reason
   Labels (User-labels)                        FlexAlgo      Metric
   Admin-tag-policy (strict-tunnel-tagging)                  Last Mod.
-------------------------------------------------------------------------------
2001:db8::2:6                                                SRV6-POLICY
   gre srv6-policy                             Y
   -- (2)                                      --            0
   -- (N)                                                    00h02m33s
```

```
-------------------------------------------------------------------------------
Next Hops : 1
===============================================================================
```

In a similar way, the BGP next hop for VPN-IPv6 routes is resolved using an SRv6 policy, as follows:

```
[/]
A:admin@PE-1# show router bgp next-hop vpn-ipv6
===============================================================================
 BGP Router ID:192.0.2.1        AS:64500       Local AS:64500
===============================================================================

===============================================================================
BGP VPN Next Hop
===============================================================================
VPN Next Hop                                                    Owner
   Autobind                                          FibProg    Reason
   Labels (User-labels)                              FlexAlgo   Metric
   Admin-tag-policy (strict-tunnel-tagging)                     Last Mod.
-------------------------------------------------------------------------------
2001:db8::2:6                                                   SRV6-POLICY
   gre srv6-policy                                   Y
   -- (2)                                            --         0
   -- (N)                                                       00h02m33s
-------------------------------------------------------------------------------
Next Hops : 1
===============================================================================
```

Conclusion

EVPN VPWS, EVPN VPLS, EVPN IFL, and VPN IP services can resolve their BGP next hops over an SRv6 policy in TTMv6.

Topology-Independent Loop-Free Alternate for Link Protection

This chapter describes the Topology-Independent Loop-Free Alternate for Link Protection.

Topics in this chapter include:

Applicability

This chapter was initially written based on SR OS Release 16.0.R5, but the MD-CLI in the current edition corresponds to SR OS Release 21.2.R1. Topology-Independent Loop-Free Alternate (TI-LFA) is supported from SR OS Release 15.0.R1 for IS-IS and 15.0.R4 for OSPF.

Overview

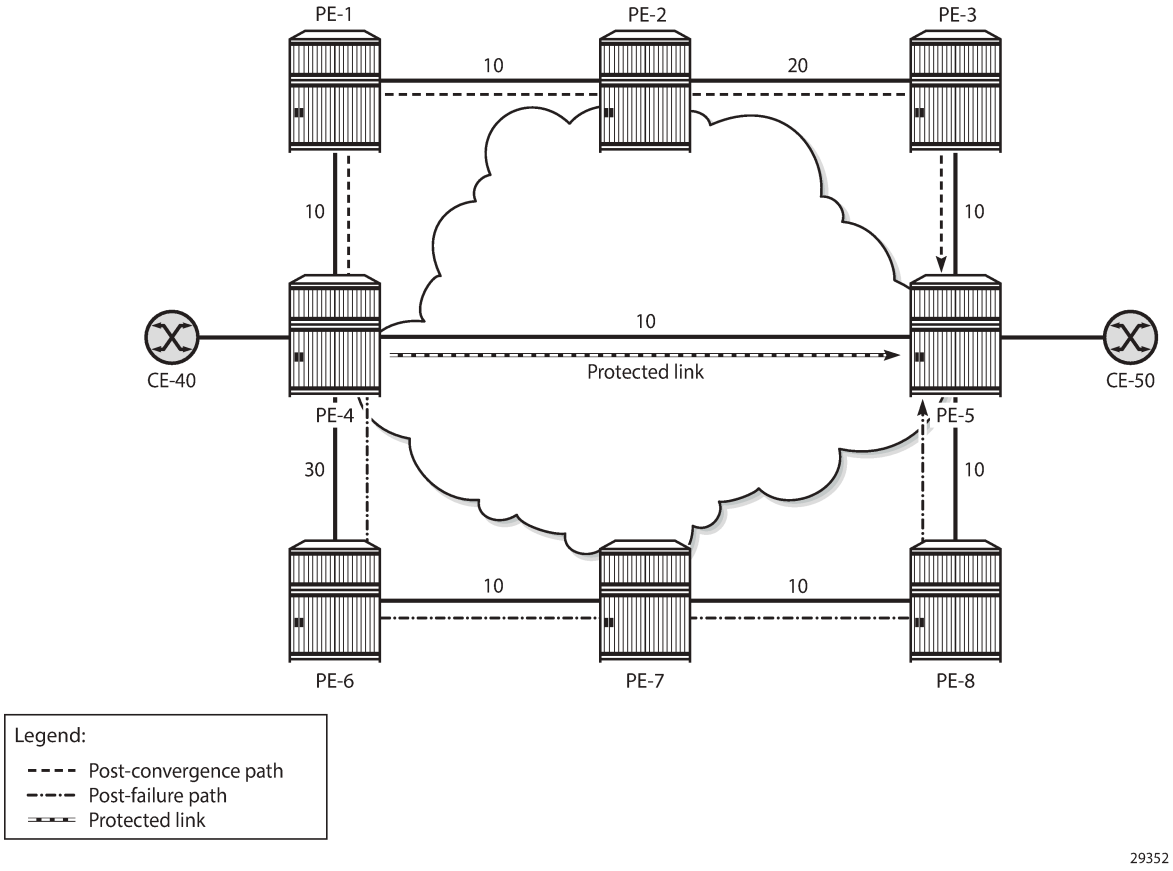
For IP Fast Reroute (FRR), the routers use a precomputed Loop-Free Alternate (LFA) next-hop installed in the FIB until the Shortest Path First (SPF) algorithm runs and the network converges again. The following LFA modes can be applied:

- Regular LFA installs an alternate next-hop in the FIB. Regular LFA provides protection for native IP traffic as well as for Segment Routing (SR) and LDP traffic.
- Remote LFA uses a repair tunnel to a PQ node, which is a node where traffic is not looped back toward the computing node. Remote LFA provides protection for SR and LDP traffic, not for native IP traffic.
- If a computing router has multiple backup next-hop routers, TI-LFA creates a repair tunnel on the post-convergence path so that the post-failure next-hop is avoided, if different from the post-convergence next-hop. In this case, traffic will not be dropped after SPF converges. TI-LFA extends the remote LFA algorithm by computing a backup tunnel where the P and Q nodes do not coincide. TI-LFA uses a repair tunnel to the closest Q node on the post-convergence path. This repair tunnel uses the shortest path to the P node and a source-routed path from the P node to the Q node. TI-LFA provides protection for SR and LDP traffic, not for native IP traffic.

Regular LFA is described in chapter MPLS LDP FRR using ISIS as IGP. Remote LFA and TI-LFA use segment routing to create repair tunnels in cases where there is no regular LFA backup.

Figure 77: Post-failure LFA path does not match post-convergence path shows the example topology where traffic flows from CE-40 toward CE-50, and a post-failure LFA path that does not match the post-convergence path.

Figure 77: Post-failure LFA path does not match post-convergence path



During normal operation, traffic goes from CE-40 to PE-4 and straight on to PE-5 and CE-50. This is the shortest path between CE-40 and CE-50. Consider the failure of the link between PE-4 and PE-5. This is the protected link. If a failure occurs on the protected link between PE-4 and PE-5, there are two possible backup next-hops from computing node PE-4: PE-1 or PE-6.

When enabling regular LFA on PE-4, two consecutive failovers will occur: the first one, nearly instantaneously, from the preferred path (optimum distance) to the precomputed post-failure path via next-hop PE-6 and the second one, after SPF has run again, from the post-failure path to the post-convergence path via PE-1. When enabling TI-LFA, a single failover will occur, so the computed post-failure path must match the post-convergence path.

The post-convergence path will be from PE-4 to PE-1, PE-2, PE-3, and PE-5, with a path cost of 10 + 10 + 20 + 10 = 50. With regular LFA, the post-failure path should not use PE-1 as next-hop, because PE-1 would loop back traffic to reach PE-5 via PE-4, through the protected link (which is not allowed).

As described in RFC 5286, the following inequality 1 for link protection must be true for a neighbor next-hop (NH) to provide an LFA. The cost is the optimum distance between the nodes:

cost(NH,Destination) < cost(NH,Source) + cost(Source,Destination)

For next-hop PE-1, the following LFA inequality 1 is false on the calculating node PE-4, indicating that no regular LFA path is possible via PE-1:

cost(PE-1,PE-5) < cost (PE-1,PE-4) + cost(PE-4,PE-5)

(10 + 10) < 10 + 10 -->*False*

For next-hop PE-6, the following LFA inequality 1 is true on the calculating node PE-4, indicating that a regular LFA path is possible via PE-6:

cost(PE-6,PE-5) < cost(PE-6,PE-4) + cost(PE-4,PE-5)

(10 + 10 + 10) < 30 + 10 -->*True*

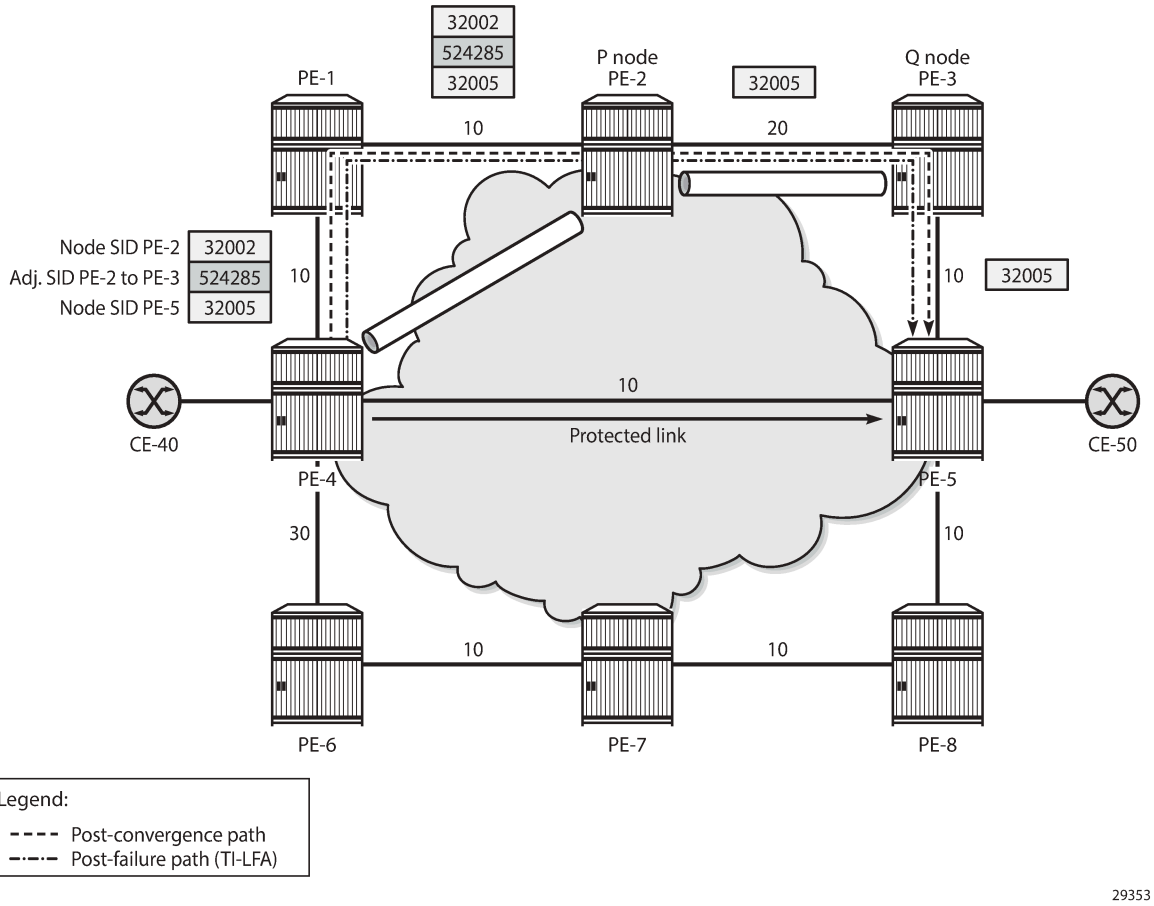
Because of the higher metric between PE-4 and PE-6 (30), PE-6 will not loop back traffic via PE-4: the path cost from PE-6 to PE-5 via PE-4 = 30 + 10 = 40, while the path cost from PE-6 to PE-5 via PE-7 and PE-8 = 10 + 10 + 10 = 30. So, PE-6 will forward the traffic to PE-7, PE-8, and PE-5.

For these reasons, the post-failure path uses PE-6 as regular LFA next-hop.

TI-LFA ensures that traffic is forwarded in a tunnel to the closest Q node, where it will not be looped back to PE-4. In this example. PE-3 is the Q node and it is one hop away from P node PE-2.

With TI-LFA enabled, additional labels are pushed to ensure that the post-failure next-hop matches the post-convergence next-hop. When the protected link between PE-4 and PE-5 fails, PE-4 pushes the node SID of PE-2 as top label plus the adjacency SID of the PE-2 to PE-3 link as an extra label. The bottom label is the node SID of the destination PE-5, which is present in any packet to PE-5 (located on the primary path); see Figure 78: Post-failure TI-LFA path matches post-convergence path.

Figure 78: Post-failure TI-LFA path matches post-convergence path



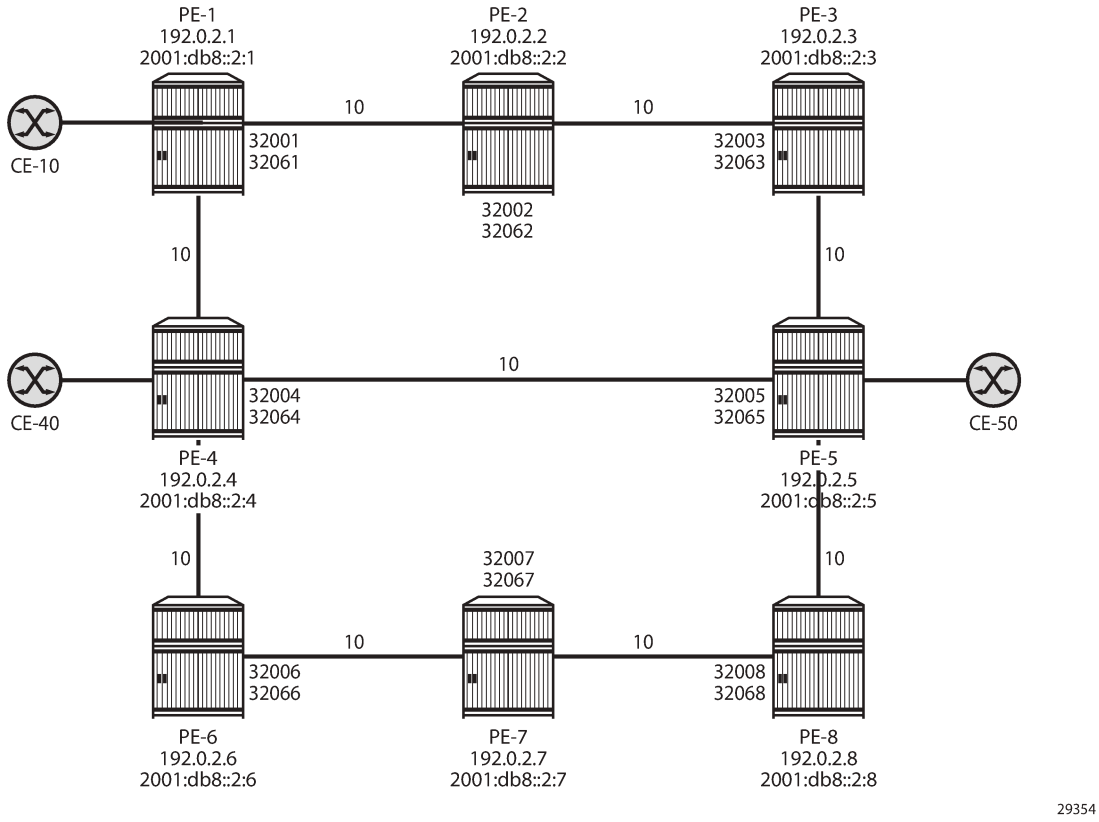
In this chapter, the following LFA modes are described and configured:

- Regular LFA
- Remote LFA
- TI-LFA

Configuration

Figure 79: Example topology shows the example topology, but that will be reduced in the first two scenarios. The default metric of all links is 10, but that may be configured with a different value afterward.

Figure 79: Example topology



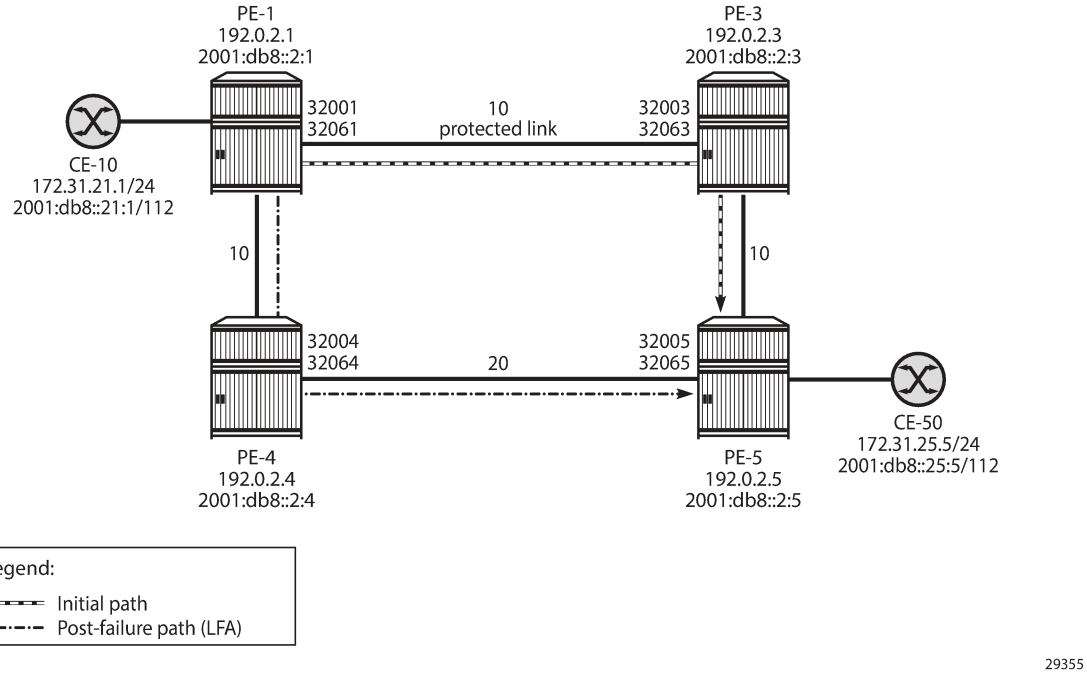
The initial configuration includes the following:

- Cards, MDAs, ports
- Dual-stack router interfaces (IPv4/IPv6)
- IS-IS as IGP on the router interfaces. The metric is 10, but that may be configured otherwise.
- Segment routing (SR-ISIS) with node SIDs 3200x for IPv4 and 3206x for IPv6 system addresses.

Regular LFA

Figure 80: Example topology with regular LFA configured on PE-4 shows the example topology reduced to four PEs. Without a failure of the protected link, traffic from CE-10 to CE-50 is sent via PE-3. The protected link is the link between PE-1 and PE-3 and the LFA path after failure goes via next-hop PE-4.

Figure 80: Example topology with regular LFA configured on PE-4



The IGP metric on the interface between PE-4 and PE-5 is 20, as follows:

```
# on PE-4:
configure {
    router "Base" {
        isis 0 {
            interface "int-PE-4-PE-5" {
                level 1 {
                    metric 20
                }
                level 2 {
                    metric 20
                } {
```

```
#on PE-5:
configure {
    router "Base" {
        isis 0 {
            interface "int-PE-5-PE-4" {
                level 1 {
                    metric 20
                }
                level 2 {
                    metric 20
                }
```

Regular LFA is configured on the nodes, as follows:

```
# on PE-1, PE-3, PE-4, PE-5:
configure {
    router "Base" {
```

```
        isis 0 {
            loopfree-alternate {
            }
```

In the normal situation, without failures, the preferred traffic path from CE-10 to CE-50 is via PE-1, PE-3, and PE-5 with a cost (optimum distance) of 10 + 10 = 20. When the link between PE-1 and PE-3 fails, the post-failure LFA path is via PE-1, PE-4, and PE-5 with a cost of 10 + 20 = 30. The following LFA inequality 1 is true, so PE-4 is a valid LFA next-hop:

cost(newNH,Destination) < cost(newNH,Source) + cost(Source,Destination)

cost(PE-4,PE-5) < cost(PE-4,PE-1) + cost(PE-1,PE-5)

20 < 10 + (10 + 10) --> True

The route table on PE-1 for prefix 192.0.2.5 shows that the next-hop is 192.168.13.2 on PE-3 for the preferred path with metric 20; the LFA next-hop is 192.168.14.2 on PE-4 for the post-failure path with metric 30, as follows:

```
[/]
A:admin@PE-1# show router route-table 192.0.2.5 alternative

===============================================================================
Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
      Alt-NextHop                                                Alt-
                                                                 Metric
-------------------------------------------------------------------------------
192.0.2.5/32                                  Remote  ISIS      00h33m04s  15
       192.168.13.2                                                20
       192.168.14.2 (LFA)                                          30
-------------------------------------------------------------------------------
No. of Routes: 1
Flags: n = Number of times nexthop is repeated
       Backup = BGP backup route
       LFA = Loop-Free Alternate nexthop
       S = Sticky ECMP requested
===============================================================================
```

The following FP tunnel table on PE-1 shows the SR-ISIS label 32005, which is the node SID of PE-5 for prefix 192.0.2.5/32. The same label 32005 is used for the LFA post-failure path indicated with (B) for FRR backup.

```
[/]
A:admin@PE-1# show router fp-tunnel-table 1 192.0.2.5/32

===============================================================================
IPv4 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                Protocol            Tunnel-ID
  Lbl
    NextHop                                                    Intf/Tunnel
  Lbl     (backup)
```

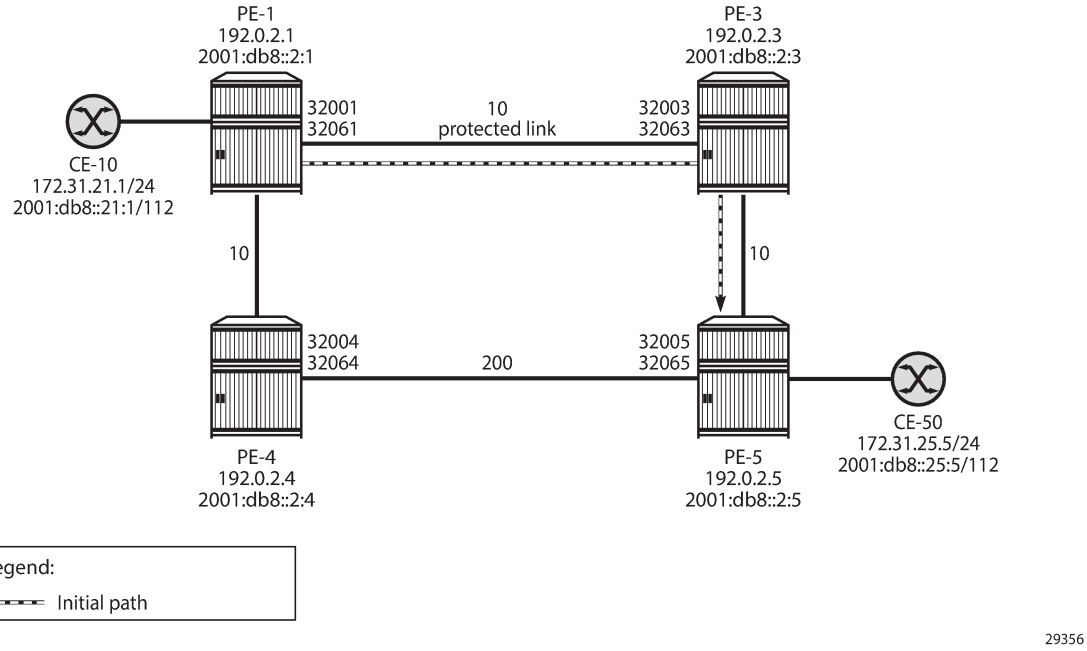
```
    NextHop   (backup)
-------------------------------------------------------------------------------
192.0.2.5/32                                SR-ISIS-0            524301
  32005
    192.168.13.2                                                 1/1/3:1000
  32005
    192.168.14.2(B)                                              1/1/2:1000
-------------------------------------------------------------------------------
Total Entries : 1
-------------------------------------------------------------------------------
===============================================================================
```

The following FP tunnel table on PE-1 shows the SR-ISIS label 32065, which is the node SID of PE-5 for prefix 2001:db8::2:5/128. The same label 32065 is used for the LFA post-failure path.

```
[/]
A:admin@PE-1# show router fp-tunnel-table 1 2001:db8::2:5/128

===============================================================================
IPv6 Tunnel Table Display

Legend:
label stack is ordered from bottom most to top-most
B - FRR Backup
===============================================================================
Destination                                 Protocol            Tunnel-ID
  Lbl
    NextHop                                                      Intf/Tunnel
  Lbl     (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
2001:db8::2:5/128                           SR-ISIS-0            524302
  32065
    fe80::618:1ff:fe01:3-"int-PE-1-PE-3"                         1/1/3:1000
  32065
    fe80::61c:1ff:fe01:1-"int-PE-1-PE-4"(B)                      1/1/2:1000
-------------------------------------------------------------------------------
Total Entries : 1
-------------------------------------------------------------------------------
===============================================================================
```

Figure 81: No post-failure LFA path when PE-4 loops back traffic shows that no backup LFA next-hop exists when the metric on the interface between PE-4 and PE-5 is increased to 200.

Figure 81: No post-failure LFA path when PE-4 loops back traffic



The following configures the metric on the interface between PE-4 and PE-5 to a value of 200:

```
# on PE-4:
configure {
    router "Base" {
        isis 0 {
            interface "int-PE-4-PE-5" {
                level 1 {
                    metric 200
                }
                level 2 {
                    metric 200
                }
```

```
# on PE-5:
configure {
    router "Base" {
        isis 0 {
            interface "int-PE-5-PE-4" {
                level 1 {
                    metric 200
                }
                level 2 {
                    metric 200
                }
```

When the metric on the interface between PE-4 and PE-5 is increased to a value that exceeds the sum of the metrics on the path from PE-4 to PE-1 and the path from PE-1 to PE-5 (via PE-3), the computing node PE-1 cannot calculate a regular LFA path to protect the PE-5 prefixes. The following LFA inequality 1 is false:

cost(PE-4,PE-5) < cost(PE-4,PE-1) + cost(PE-1,PE-5)

200 < 10 + (10 + 10) --> False

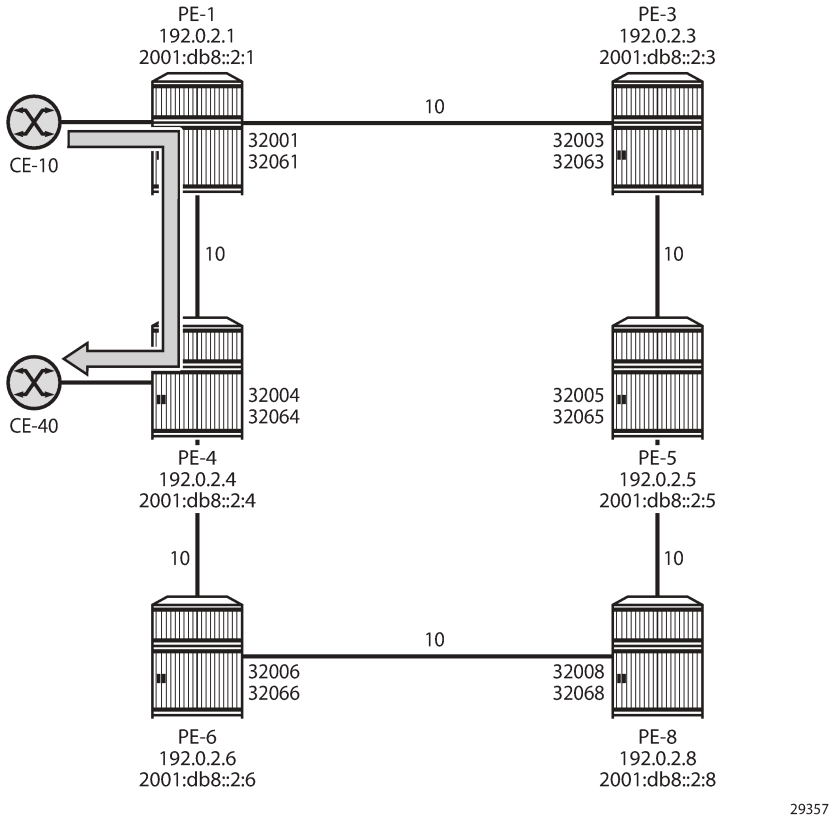
If the preferred path cannot be used because of a failure, such as a link failure between PE-1 and PE-3, a micro-loop is created between PE-4 and PE-5 until convergence is completed. The following output shows that no LFA next-hop is available on PE-1:

```
[/]
A:admin@PE-1# show router route-table 192.0.2.5 alternative

===============================================================================
Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
      Alt-NextHop                                                Alt-
                                                                 Metric
-------------------------------------------------------------------------------
192.0.2.5/32                                  Remote  ISIS      00h05m26s  15
       192.168.13.2                                                 20
-------------------------------------------------------------------------------
No. of Routes: 1
Flags: n = Number of times nexthop is repeated
       Backup = BGP backup route
       LFA = Loop-Free Alternate nexthop
       S = Sticky ECMP requested
===============================================================================
```

Remote LFA

Figure 82: Example topology for remote LFA shows the example topology with six nodes in a ring. Traffic from CE-10 to CE-40 is preferably sent via PE-1 to PE-4.

Figure 82: Example topology for remote LFA



The following command enables remote LFA on all nodes:

```
# on PE-1, PE-3, PE-4, PE-5, PE-6, PE-8:
configure {
    router "Base" {
        isis 0 {
            loopfree-alternate {
                remote-lfa {
                }
            }
```

If the link between PE-1 and PE-4 fails, the repair path on PE-1 can only use PE-3 as next-hop. Link-protection LFA inequality 1 is not valid, indicating that the backup path via next-hop PE-3 is not loop free:

cost(PE-3,PE-4) < cost(PE-3,PE-1) + cost(PE-1,PE-4)

(10 + 10) < 10 + 10 --> *False*

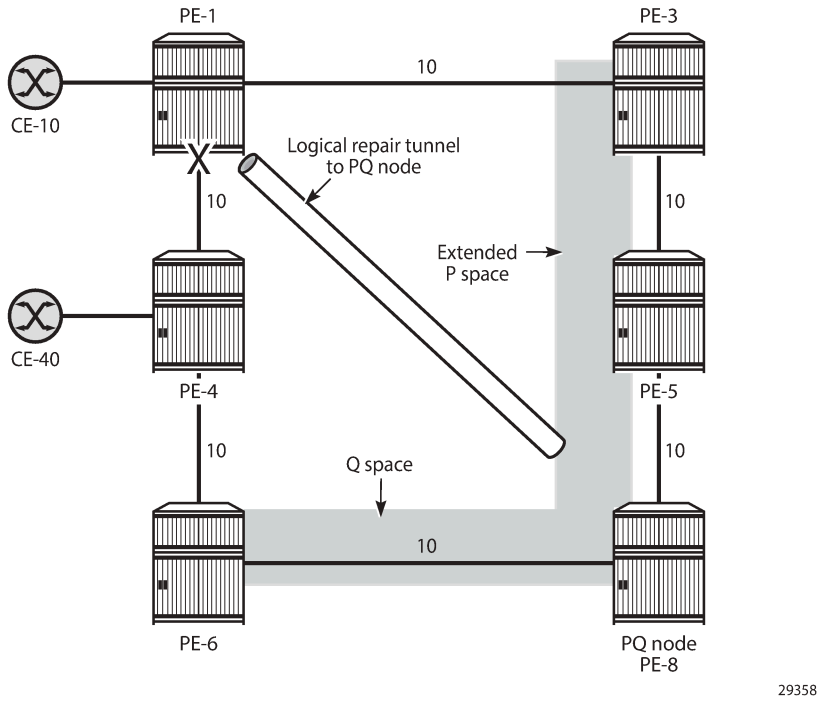
With this invalid LFA inequality 1, no coverage with regular LFA is possible. When remote LFA is enabled, a repair tunnel is computed from PE-1 toward a node (PE-8) where the traffic is not looped back toward the computing node PE-1. When traffic emerges from the repair tunnel on PE-8, it is forwarded to the destination PE-4, using node SID 32004 for IPv4 or node SID 32064 for IPv6.

The endpoint node of the repair tunnel for remote LFA (RLFA) is the PQ node, which is in the intersection of the extended P space of source PE-1 and the Q space of destination PE-4.

- The P space of PE-1 is the set of routers reachable on the shortest SPF path from the computing node PE-1, without using the protected link between PE-1 and PE-4; that is, SPF computed by PE-1 and rooted from PE-1. In this example, PE-3 and PE-5 are in the P space of PE-1.
- The extended P space of PE-1 is the set of routers, calculated by PE-1, in the P space of the next-hop router PE-3. An additional SPF computation by PE-1 and rooted from PE-3 results in P nodes PE-3, PE-5, and PE-8. The extended P space increases the repair coverage.
- The Q space of PE-4 is the set of routers that can reach the destination router PE-4 using the shortest path, without using the protected link; that is, reverse SPF computed by PE-1 and rooted from PE-4, resulting in Q nodes PE-6 and PE-8.
- PQ routers are in the intersection of the extended P space and the Q space; in this case, the only PQ node is PE-8.
- Repair tunnels are shortest path SR tunnels from the computing node PE-1 to the PQ router; in this case, from PE-1 to PE-8.

Figure 83: PQ node in remote LFA shows the extended P space of PE-1, comprising nodes PE-3, PE-5, and PE-8, and the Q space of PE-4, comprising nodes PE-6 and PE-8. In the event of a link failure, PE-1 will push the node SID of PE-8, along with the node SID of PE-4, and forward the packet toward the backup next-hop PE-8.

Figure 83: PQ node in remote LFA



The following shows the SR LFA coverage on PE-1; the five other node SIDS are all protected: one with regular LFA and the remaining four with remote LFA (in the column RLFA). Besides the node SIDs, the

adjacency SIDs toward the direct neighbors PE-3 and PE-4 are protected using RLFA. The LFA coverage is the same for IPv4 and IPv6.

```
[/]
A:admin@PE-1# show router isis sr-lfa-coverage

===============================================================================
Rtr Base ISIS Instance 0 SR LFA Coverage
===============================================================================
MT-ID   SidType      Level Proto LFA       RLFA      TILFA     Coverage
-------------------------------------------------------------------------------
0       node-sid     L1    ipv4  1(20%)    4(80%)    0(0%)     5/5(100%)
0       node-sid     L1    ipv6  1(20%)    4(80%)    0(0%)     5/5(100%)
---snip---
0       adj-sid      L1L2  ipv4  0(0%)     2(100%)   0(0%)     2/2(100%)
0       adj-sid      L1L2  ipv6  0(0%)     2(100%)   0(0%)     2/2(100%)
===============================================================================
```

The repair tunnel from PE-1 to PQ node PE-8 uses node SID 32008 for IPv4 and 32068 for IPv6.

The fifth entry in the following FP tunnel table shows that destination 192.0.2.8/32 of PE-8 is protected with regular LFA. The only label is 32008, which is the node SID of PE-8. All other destinations in the table are protected with remote LFA, having two node SID labels for the RLFA path, such as 32004/32008 for prefix 192.0.2.4 with next-hop 192.168.13.2 on PE-3. This means that the top label 32008 is pushed by PE-1 to match the repair-tunnel going via PE-3 to PQ-node PE-8. From PE-8 onward, the bottom label 32004 is used toward PE-4. Likewise, the other destinations in the list have top label 32008, so a tunnel is established to PE-8. The output is similar for IPv6.

```
[/]
A:admin@PE-1# show router fp-tunnel-table 1

===============================================================================
IPv4 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                Protocol            Tunnel-ID
  Lbl
    NextHop                                                    Intf/Tunnel
  Lbl     (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
192.0.2.3/32                               SR-ISIS-0            524295
  32003
    192.168.13.2                                               1/1/3:1000
  32003/32008
    192.168.14.2(B)                                            1/1/2:1000
192.0.2.4/32                               SR-ISIS-0            524299
  32004
    192.168.14.2                                               1/1/2:1000
  32004/32008
    192.168.13.2(B)                                            1/1/3:1000
192.0.2.5/32                               SR-ISIS-0            524301
  32005
    192.168.13.2                                               1/1/3:1000
  32005/32008
    192.168.14.2(B)                                            1/1/2:1000
192.0.2.6/32                               SR-ISIS-0            524311
  32006
```

```
    192.168.14.2                                                1/1/2:1000
  32006/32008
    192.168.13.2(B)                                             1/1/3:1000
192.0.2.8/32                            SR-ISIS-0               524312
  32008
    192.168.13.2                                                1/1/3:1000
  32008
    192.168.14.2(B)                                             1/1/2:1000
192.168.13.2/32                         SR                      524309
  3
    192.168.13.2                                                1/1/3:1000
  32003/32008
    192.168.14.2(B)                                             1/1/2:1000
192.168.14.2/32                         SR                      524297
  3
    192.168.14.2                                                1/1/2:1000
  32004/32008
    192.168.13.2(B)                                             1/1/3:1000
-------------------------------------------------------------------------------
Total Entries : 7
-------------------------------------------------------------------------------
===============================================================================
```

TI-LFA

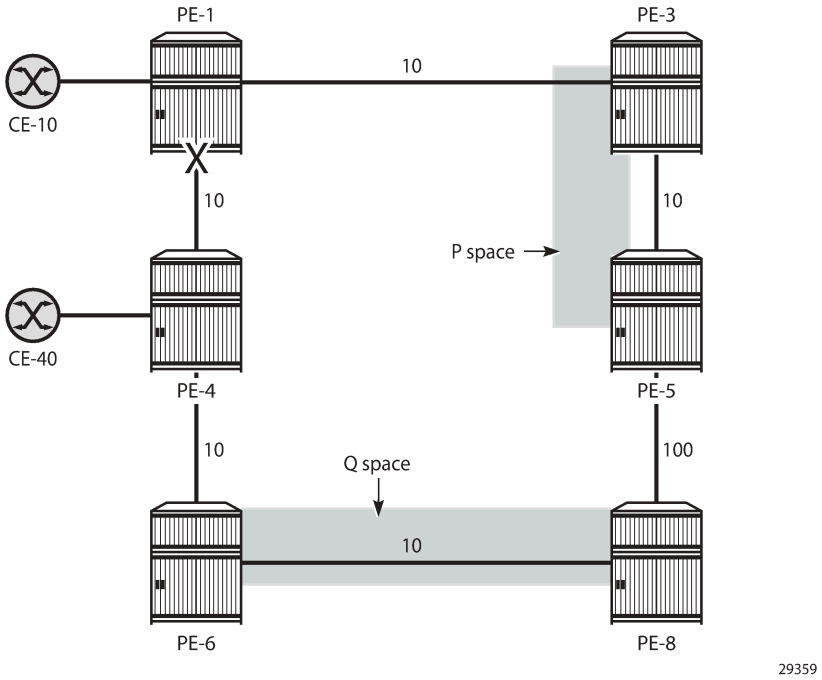
The following two use cases are described in this section:

- Directed LFA where the extended P space and the Q space do not overlap
- Extension of the RLFA algorithm to compute a repair path using directed LFA, but ensuring that the post-failure path matches the post-convergence path

Directed LFA

Figure 84: Extended P space of PE-1 and Q space of PE-4 are one hop apart shows the example topology with increased metric between PE-5 and PE-8, reducing the extended P space to PE-3 and PE-5, so there is no PQ node.

Figure 84: Extended P space of PE-1 and Q space of PE-4 are one hop apart



There is no remote LFA repair tunnel. No Q routers are on the shortest path from the computing router, and the P routers are not in the reverse SPF of the endpoint of the protected link. However, TI-LFA can calculate a repair tunnel in case the gap is only one or two hops. TI-LFA is enabled using the following command:

```
# on PE-1, PE-3, PE-4, PE-5, PE-6, PE-8:
configure {
    router "Base" {
        isis 0 {
            loopfree-alternate {
                ti-lfa {
                    max-sr-frr-labels 2
                }
```

Table 9: Values of the max-sr-frr-labels parameter in TI-LFA lists the possible values of the **max-sr-frr-labels** parameter. This parameter is used to specify the maximum number of labels that the TI-LFA backup next-hop can use.

Table 9: Values of the max-sr-frr-labels parameter in TI-LFA

Max. SR-FRR labels	LFA behavior
0	Regular LFA: TI-LFA backup restricted to next-hop that does not require a repair tunnel, so PQ node is a neighbor of the computing node.
1	Remote LFA: extended P space and Q space intersect and the repair tunnel requires 1 FRR label:

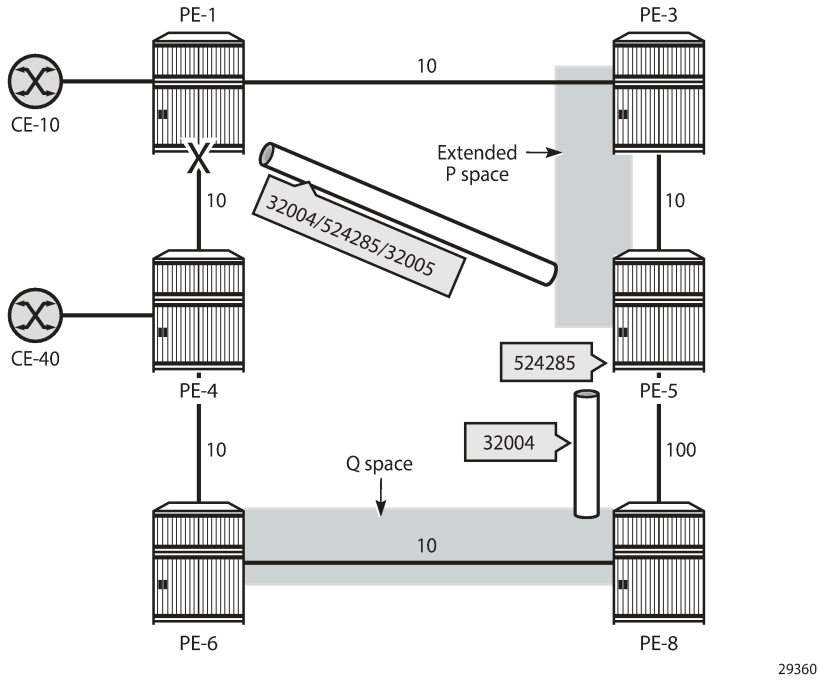
Max. SR-FRR labels	LFA behavior
	• Node SID to PQ router
2 (default)	TI-LFA with extended P space and Q space one hop apart: • Node SID to P router • Adjacency SID on P router to Q router
3	TI-LFA with extended P space and Q space two hops apart: • Node SID to P router • Two adjacency SIDs to Q router

In this case, the extended P space and the Q space are one hop apart and TI-LFA calculates a post-failure path that consists of a repair tunnel to P router PE-5 (node SID 32005 for IPv4) and an adjacency SID toward Q router PE-8. For routes from PE-1 to PE-4, the LFA route has two additional labels combined with the bottom label that is the node SID of PE-4 (32004), which is also used for the primary path. The top label is the node SID of P router PE-5 (32005); the next label is the adjacency SID on PE-5 toward PE-8 (524285).

```
[/]
A:admin@PE-1# show router fp-tunnel-table 1 192.0.2.4/32

===============================================================================
IPv4 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                Protocol            Tunnel-ID
  Lbl
    NextHop                                                    Intf/Tunnel
  Lbl     (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
192.0.2.4/32                               SR-ISIS-0            524299
  32004
    192.168.14.2                                               1/1/2:1000
  32004/524285/32005
    192.168.13.2(B)                                            1/1/3:1000
-------------------------------------------------------------------------------
Total Entries : 1
-------------------------------------------------------------------------------
===============================================================================
```

Figure 85: Directed LFA with P router and Q router one hop apart shows the directed LFA path from source PE-1 to P router PE-5 (node SID), the adjacency SID from P router PE-5 to Q router PE-8, and the node SID of destination PE-4. P router PE-5 uses the adjacency SID for forwarding, but only sends the packets with the node SID of PE-4 (32004).

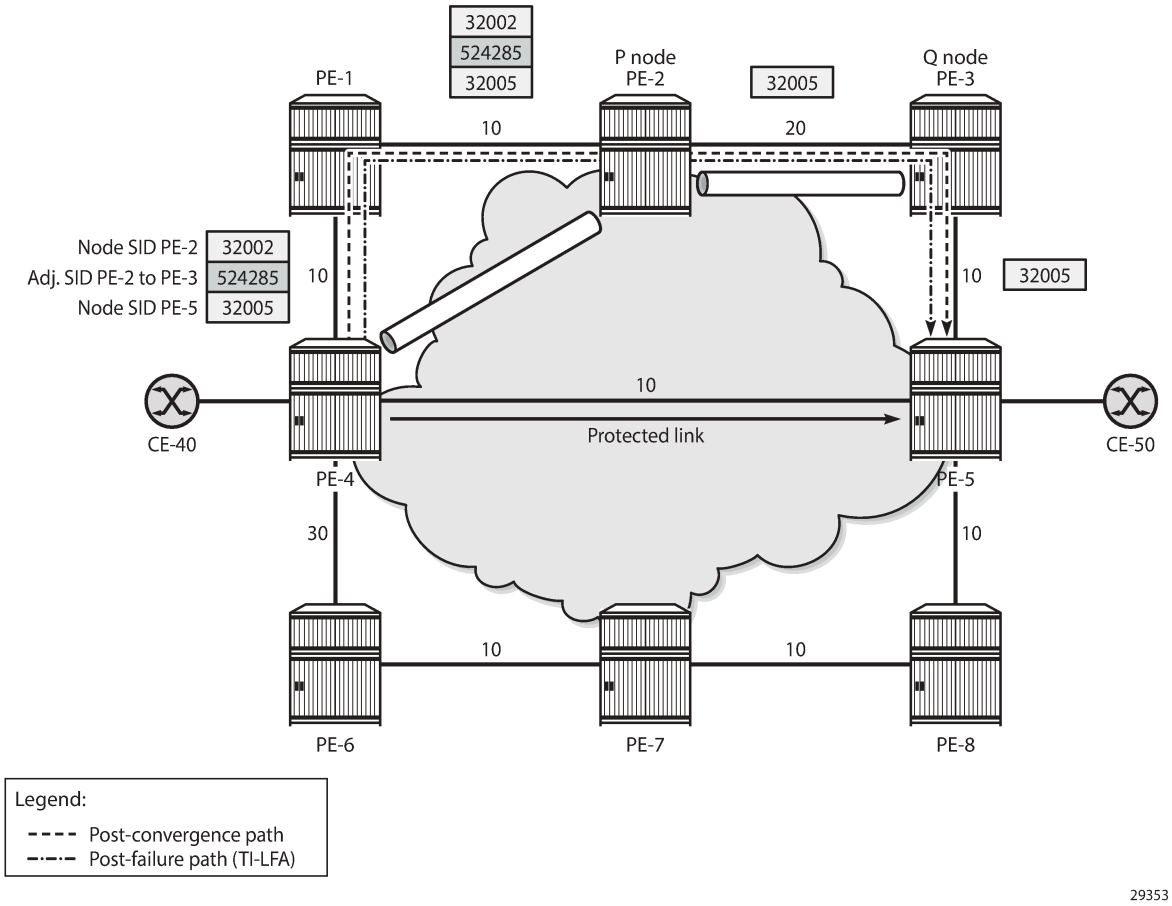
Figure 85: Directed LFA with P router and Q router one hop apart



TI-LFA for coinciding post-failure and post-convergence paths

Figure 86: Post-failure TI-LFA path coincides with post-convergence path is the same as Figure 78: Post-failure TI-LFA path matches post-convergence path and is repeated here for readability. The router interfaces have IGP metric 10 by default, except for the interfaces between PE-2 and PE-3 that have metric 20, and the interfaces between PE-4 and PE-6 that have metric 30. As in Figure 85: Directed LFA with P router and Q router one hop apart, Figure 86: Post-failure TI-LFA path coincides with post-convergence path shows the different tunnels used for the TI-LFA path. TI-LFA ensures that the post-failure path coincides with the post-convergence path by adding additional labels: the node SID 32002 (or 32062 for IPv6) to P router PE-2, the adjacency SID on PE-2 for the interface toward Q router PE-3, and the node SID 32005 (or 32065 for IPv6) toward the destination PE-5.

Figure 86: Post-failure TI-LFA path coincides with post-convergence path



Regular LFA coverage

For a better comparison, the regular LFA coverage is calculated first. Without remote LFA and TI-LFA enabled, the LFA coverage is limited. The following command disables remote LFA and TI-LFA on all nodes, while regular LFA remains enabled:

```
# on all nodes:
configure {
    router "Base" {
        isis 0 {
            loopfree-alternate {
                delete remote-lfa
                delete ti-lfa
            }
```

The SR LFA coverage on PE-4 only protects node SIDs and adjacency SIDs that can be protected with regular LFA, as follows:

```
[/]
```

```
A:admin@PE-4# show router isis sr-lfa-coverage

===============================================================================
Rtr Base ISIS Instance 0 SR LFA Coverage
===============================================================================
MT-ID   SidType      Level Proto LFA       RLFA      TILFA     Coverage
-------------------------------------------------------------------------------
0       node-sid     L1    ipv4  5(71%)    0(0%)     0(0%)     5/7(71%)
0       node-sid     L1    ipv6  5(71%)    0(0%)     0(0%)     5/7(71%)
---snip---
0       adj-sid      L1L2  ipv4  2(66%)    0(0%)     0(0%)     2/3(66%)
0       adj-sid      L1L2  ipv6  2(66%)    0(0%)     0(0%)     2/3(66%)
===============================================================================
```

The following shows that no LFA paths exist on PE-4 for destinations 192.0.2.1 (PE-1), 192.0.2.2 (PE-2), and 192.168.14.1 (PE-1).

```
[/]
A:admin@PE-4# show router fp-tunnel-table 1

===============================================================================
IPv4 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                Protocol            Tunnel-ID
  Lbl
    NextHop                                                    Intf/Tunnel
  Lbl     (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
192.0.2.1/32                               SR-ISIS-0            524291
  32001
    192.168.14.1                                               1/1/1:1000
192.0.2.2/32                               SR-ISIS-0            524319
  32002
    192.168.14.1                                               1/1/1:1000
192.0.2.3/32                               SR-ISIS-0            524293
  32003
    192.168.45.2                                               1/1/3:1000
  32003
    192.168.46.2(B)                                            1/1/2:1000
192.0.2.5/32                               SR-ISIS-0            524299
  32005
    192.168.45.2                                               1/1/3:1000
  32005
    192.168.46.2(B)                                            1/1/2:1000
192.0.2.6/32                               SR-ISIS-0            524311
  32006
    192.168.46.2                                               1/1/2:1000
  32006
    192.168.45.2(B)                                            1/1/3:1000
192.0.2.7/32                               SR-ISIS-0            524323
  32007
    192.168.45.2                                               1/1/3:1000
  32007
    192.168.46.2(B)                                            1/1/2:1000
192.0.2.8/32                               SR-ISIS-0            524313
  32008
    192.168.45.2                                               1/1/3:1000
  32008
```

```
    192.168.46.2(B)                                                 1/1/2:1000
192.168.14.1/32                          SR                         524317
  3
    192.168.14.1                                                    1/1/1:1000
192.168.45.2/32                          SR                         524321
  3
    192.168.45.2                                                    1/1/3:1000
  32005
    192.168.46.2(B)                                                 1/1/2:1000
192.168.46.2/32                          SR                         524309
  3
    192.168.46.2                                                    1/1/2:1000
  32006
    192.168.45.2(B)                                                 1/1/3:1000
-------------------------------------------------------------------------------
Total Entries : 10
-------------------------------------------------------------------------------
===============================================================================
```

For destination 192.0.2.5, the post-failure path has next-hop 192.168.46.2 on PE-6, so the post-failure path does not coincide with the post-convergence path with next-hop 192.168.14.1 on PE-1. The path cost of the post-convergence path from PE-4 to PE-5 (via PE-1, PE-2, and PE-3) equals 10 + 10 + 20 + 10 = 50; the path cost of the post-failure path from PE-4 to PE-5 (via PE-6, PE-7, and PE-8) equals 30 + 10 + 10 + 10 = 60.

TI-LFA enabled

TI-LFA can be configured with remote LFA enabled or disabled. The following command configures remote LFA and TI-LFA (with default max-sr-frr-labels 2).

```
# on all nodes:
configure {
    router "Base" {
        isis 0 {
            loopfree-alternate {
                remote-lfa {
                }
                ti-lfa {
                    max-sr-frr-labels 2
                }
```

With TI-LFA enabled, the SR LFA coverage increases to 100%, as follows. For almost all destinations, the LFA protection is now using TI-LFA, even when regular LFA was possible before. The advantage is that TI-LFA ensures the post-failure path coincides with the post-convergence path.

If there is regular LFA protection via a path that does not coincide with the post-convergence path, that regular LFA protection will only change to TI-LFA protection when max-sr-frr-labels allows the needed number of labels (tunnels) to force the TI-LFA protection to the post-convergence path. The same applies for remote LFA protection.

```
[/]
A:admin@PE-4# show router isis sr-lfa-coverage

===============================================================================
Rtr Base ISIS Instance 0 SR LFA Coverage
===============================================================================
MT-ID   SidType      Level Proto LFA       RLFA      TILFA     Coverage
-------------------------------------------------------------------------------
```

```
0       node-sid     L1    ipv4  0(0%)     0(0%)     7(100%)   7/7(100%)
0       node-sid     L1    ipv6  0(0%)     0(0%)     7(100%)   7/7(100%)
---snip---
0       adj-sid      L1L2  ipv4  0(0%)     0(0%)     3(100%)   3/3(100%)
0       adj-sid      L1L2  ipv6  0(0%)     0(0%)     3(100%)   3/3(100%)
===============================================================================
```

The following FP tunnel table shows that prefixes 192.0.2.1 (PE-1), 192.0.2.2 (PE-2), and 192.168.14.1 (PE-1) are now protected too. For destination 192.0.2.5 (PE-5), the next-hop now is 192.168.14.1, which is also the next-hop on the post-convergence path to PE-5 via PE-1, PE-2, and PE-3. The top label 32002 is the node SID of PE-2, the label 524285 is the adjacency SID on PE-2 for the interface toward PE-3, and the bottom label 32005 is the node SID to reach the destination PE-5.

```
[/]
A:admin@PE-4# show router fp-tunnel-table 1

===============================================================================
IPv4 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                Protocol            Tunnel-ID
  Lbl
    NextHop                                                    Intf/Tunnel
  Lbl     (backup)
    NextHop   (backup)
-------------------------------------------------------------------------------
192.0.2.1/32                               SR-ISIS-0           524291
  32001
    192.168.14.1                                               1/1/1:1000
  32001/524285/32003
    192.168.45.2(B)                                            1/1/3:1000
192.0.2.2/32                               SR-ISIS-0           524319
  32002
    192.168.14.1                                               1/1/1:1000
  32002/524285/32003
    192.168.45.2(B)                                            1/1/3:1000
192.0.2.3/32                               SR-ISIS-0           524293
  32003
    192.168.45.2                                               1/1/3:1000
  32003/524285/32002
    192.168.14.1(B)                                            1/1/1:1000
192.0.2.5/32                               SR-ISIS-0           524299
  32005
    192.168.45.2                                               1/1/3:1000
  32005/524285/32002
    192.168.14.1(B)                                            1/1/1:1000
192.0.2.6/32                               SR-ISIS-0           524311
  32006
    192.168.46.2                                               1/1/2:1000
  32006
    192.168.45.2(B)                                            1/1/3:1000
192.0.2.7/32                               SR-ISIS-0           524323
  32007
    192.168.45.2                                               1/1/3:1000
  32007
    192.168.46.2(B)                                            1/1/2:1000
192.0.2.8/32                               SR-ISIS-0           524313
  32008
    192.168.45.2                                               1/1/3:1000
```

```
  32008
    192.168.46.2(B)                                                 1/1/2:1000
192.168.14.1/32                          SR                         524317
  3
    192.168.14.1                                                    1/1/1:1000
  32001/524285/32003
    192.168.45.2(B)                                                 1/1/3:1000
192.168.45.2/32                          SR                         524321
  3
    192.168.45.2                                                    1/1/3:1000
  32005/524285/32002
    192.168.14.1(B)                                                 1/1/1:1000
192.168.46.2/32                          SR                         524309
  3
    192.168.46.2                                                    1/1/2:1000
  32006
    192.168.45.2(B)                                                 1/1/3:1000
-------------------------------------------------------------------------------
Total Entries : 10
-------------------------------------------------------------------------------
===============================================================================
```

The following **tools** command on PE-4 includes detailed information for the LFA protection for destination 192.0.2.5:

```
[/]
A:admin@PE-4# tools dump router isis sr-database prefix 192.0.2.5 detail
===============================================================================
Rtr Base ISIS Instance 0 SR Database

Legend:
label stack is ordered from bottom-most to top-most
===============================================================================
-------------------------------------------------------------------------------
SID 5
-------------------------------------------------------------------------------
Label              : 32005                Adv System Id       : 1920.0000.2005
Prefix             : 192.0.2.5
Route Level        : 1                    MT Id               : 0
Rtm Preference     : 15                   Ttm Preference      : 11
Metric             : 10                   Last Action         : LfaNhops
Num Ip NextHop     : 1                    Num SR-Tnl NextHop  : 1
Mtu                : 8970
Mtu Prim           : 8982                 Mtu Backup          : 8982
Exclude from LFA   : 0                    LFA Type            : TI LFA
Duplicate Pending  : 0                    Tunnel Active State : Reported/Ack
SR Error           : SR_ERR_OK

LFA NextHop IP     : 192.168.14.1
LFA IsTunl         : N
LFA GIfId/TunlType : 1                    LFA IfId/LspId      : 2
LFA PgId           : 0                    LFA Adv Node        : False
LFA Labels         : 32005/524285/32002

NHOP: IP                       IsTunl GIfId/   IfId/ PgId  IsAdv Label   IsLfaX
                                      TunlType LspId
-------------------------------------------------------------------------------
192.168.45.2                   N      2        3     13    1     32005   0

-------------------------------------------------------------------------------
No. of Entries: 1
-------------------------------------------------------------------------------
LDP = LDP FEC is the SID NH for SR-LDP stitching
```

```
===============================================================================
```

TI-LFA enabled with max-sr-frr-labels lower than 2

When TI-LFA is configured with max-sr-frr-labels lower than 2, TI-LFA cannot substitute regular or remote LFA where more than 2 tunnel labels are needed for the substitution. Some destinations may remain protected then via regular or remote LFA, and only those destinations that can be protected with TI-LFA with less than 2 tunnel labels will have TI-LFA protection. The following configuration enables TI-LFA with max-sr-frr-labels equal to 1:

```
# on all nodes:
configure {
    router "Base" {
        isis 0 {
            loopfree-alternate {
                remote-lfa {
                }
                ti-lfa {
                    max-sr-frr-labels 1
                }
```

In the topology of Figure 86: Post-failure TI-LFA path coincides with post-convergence path, for max-sr-frr-labels equal to 1, the SR LFA coverage drops below 100% again, as follows.

```
[/]
A:admin@PE-4# show router isis sr-lfa-coverage

===============================================================================
Rtr Base ISIS Instance 0 SR LFA Coverage
===============================================================================
MT-ID   SidType      Level Proto LFA       RLFA      TILFA     Coverage
-------------------------------------------------------------------------------
0       node-sid     L1    ipv4  2(28%)    0(0%)     3(42%)    5/7(71%)
0       node-sid     L1    ipv6  2(28%)    0(0%)     3(42%)    5/7(71%)
---snip---
0       adj-sid      L1L2  ipv4  1(33%)    0(0%)     1(33%)    2/3(66%)
0       adj-sid      L1L2  ipv6  1(33%)    0(0%)     1(33%)    2/3(66%)
===============================================================================
```

The preceding information can be derived from the FP tunnel table and the SR database as follows. For PE-4, the FP tunnel table shows that there are 10 destinations, 7 nodes and 3 next-hops. 5 out of 7 node destinations and 2 out of 3 next-hop destinations are protected with a backup (B). Node destination 192.0.2.1 (PE-1), and 192.0.2.2 (PE-2), and next-hop destination 192.168.14.1 are no longer protected.

```
[/]
A:admin@PE-4# show router fp-tunnel-table 1

===============================================================================
IPv4 Tunnel Table Display

Legend:
label stack is ordered from bottom-most to top-most
B - FRR Backup
===============================================================================
Destination                                 Protocol            Tunnel-ID
  Lbl
    NextHop                                                     Intf/Tunnel
  Lbl     (backup)
```

```
    NextHop   (backup)
-------------------------------------------------------------------------------
192.0.2.1/32                                SR-ISIS-0           524291
  32001
    192.168.14.1                                                1/1/1:1000
192.0.2.2/32                                SR-ISIS-0           524319
  32002
    192.168.14.1                                                1/1/1:1000
192.0.2.3/32                                SR-ISIS-0           524293
  32003
    192.168.45.2                                                1/1/3:1000
  32003
    192.168.46.2(B)                                             1/1/2:1000
192.0.2.5/32                                SR-ISIS-0           524299
  32005
    192.168.45.2                                                1/1/3:1000
  32005
    192.168.46.2(B)                                             1/1/2:1000
192.0.2.6/32                                SR-ISIS-0           524311
  32006
    192.168.46.2                                                1/1/2:1000
  32006
    192.168.45.2(B)                                             1/1/3:1000
192.0.2.7/32                                SR-ISIS-0           524323
  32007
    192.168.45.2                                                1/1/3:1000
  32007
    192.168.46.2(B)                                             1/1/2:1000
192.0.2.8/32                                SR-ISIS-0           524313
  32008
    192.168.45.2                                                1/1/3:1000
  32008
    192.168.46.2(B)                                             1/1/2:1000
192.168.14.1/32                             SR                  524317
  3
    192.168.14.1                                                1/1/1:1000
192.168.45.2/32                             SR                  524325
  3
    192.168.45.2                                                1/1/3:1000
  32005
    192.168.46.2(B)                                             1/1/2:1000
192.168.46.2/32                             SR                  524309
  3
    192.168.46.2                                                1/1/2:1000
  32006
    192.168.45.2(B)                                             1/1/3:1000
-------------------------------------------------------------------------------
Total Entries : 10
-------------------------------------------------------------------------------
===============================================================================
```

The SR database indicates what type of protection corresponds with the (topmost) label of the destinations in the FP tunnel tabel. Destination 192.0.2.1 (PE-1), 192.0.2.2 (PE-2), and 192.168.14.1 have no backup. Their label indicates that there is no LFA protection (LT = –). Destination 192.0.2.3 (PE-3), 192.0.2.5 (PE-5), and 192.168.45.2 have a backup with a (topmost) label that indicates regular LFA protection (LT = L). So, destination 192.0.2.5 (PE-5) is no longer TI-LFA protected, because that would require 2 tunnel labels, which max-sr-frr-labels=1 prevents. Destination 192.0.2.6 (PE-6), 192.0.2.7(PE-7), 192.0.2.8 (PE-8), and 192.168.46.2 have a backup with a (topmost) label that indicates TI-LFA protection (LT = T). As these destinations have no TI-LFA tunnel label, their TI-LFA protection does not need tunnels to ensure that the TI-LFA protection is via the post-convergence path.

The following **tools** command on PE-4 includes detailed information for the type of LFA protection that corresponds with a label:

```
[/]
A:admin@PE-4# tools dump router isis sr-database ipv4-unicast
===============================================================================
Rtr Base ISIS Instance 0 SR Database
===============================================================================
SID   Label  Prefix          Last-act  Lev MT RtmPref TtmPref Metric  IpNh SrNh
  Mtu    MtuPrim MtuBk  D xL LT Act AdvSystemId    SrErr
-------------------------------------------------------------------------------
1     32001  192.0.2.1       RemLfaNh  1   0  15      11      10      1    1
  8974   8982    -      0 0  -  +R  1920.0000.2001 SR_ERR_OK

2     32002  192.0.2.2       RemLfaNh  1   0  15      11      20      1    1
  8974   8982    -      0 0  -  +R  1920.0000.2002 SR_ERR_OK

3     32003  192.0.2.3       LfaNhops  1   0  15      11      20      1    1
  8974   8982    8982   0 0  L  +R  1920.0000.2003 SR_ERR_OK

4     32004  192.0.2.4       Local     -   -  -       -       -       -    -
  -      -       -      0 -  -  +R  1920.0000.2004 SR_ERR_OK

5     32005  192.0.2.5       LfaNhops  1   0  15      11      10      1    1
  8974   8982    8982   0 0  L  +R  1920.0000.2005 SR_ERR_OK

6     32006  192.0.2.6       TnlChange 1   0  15      11      30      1    1
  8974   8982    8982   0 0  T  +R  1920.0000.2006 SR_ERR_OK

7     32007  192.0.2.7       TnlChange 1   0  15      11      30      1    1
  8974   8982    8982   0 0  T  +R  1920.0000.2007 SR_ERR_OK

8     32008  192.0.2.8       TnlChange 1   0  15      11      20      1    1
  8974   8982    8982   0 0  T  +R  1920.0000.2008 SR_ERR_OK

-------------------------------------------------------------------------------
No. of Entries: 8
-------------------------------------------------------------------------------
Lev = route level
IpNh = number of IP next-hops
SrNh = number of SR-tunnel next-hops
D = duplicate pending
xL = exclude from LFA
LT = LFA type (L:LFA, R:RLFA, T:TILFA, n:nodeProtection)
Act = tunnel active state (R:reported, F:failed, +:SR-ack)
===============================================================================
```

Independent from the preceding ISIS Segment Routing LFA coverage (per Segment Routing LFA type and per ISIS Level), there is also the ISIS IP-routing LFA coverage (per IP version and per ISIS Level), as follows:

```
[/]
A:admin@PE-4# show router isis lfa-coverage

===============================================================================
Rtr Base ISIS Instance 0 LFA Coverage
===============================================================================
Topology         Level   Node           IPv4                IPv6
-------------------------------------------------------------------------------
IPV4 Unicast     L1      5/7(71%)       9/13(69%)           9/13(69%)
IPV6 Unicast     L1      0/0(0%)        0/0(0%)             0/0(0%)
IPV4 Multicast   L1      0/0(0%)        0/0(0%)             0/0(0%)
```

```
IPV6 Multicast   L1      0/0(0%)        0/0(0%)            0/0(0%)
IPV4 Unicast     L2      5/7(71%)       9/13(69%)          9/13(69%)
IPV6 Unicast     L2      0/0(0%)        0/0(0%)            0/0(0%)
IPV4 Multicast   L2      0/0(0%)        0/0(0%)            0/0(0%)
IPV6 Multicast   L2      0/0(0%)        0/0(0%)            0/0(0%)
===============================================================================
```

The preceding information can be derived from the table of alternative ISIS routes as follows. For PE-4, there are 17 routes: 8 routes to nodes and 9 routes to networks. The node and the networks that have 0.0.0.0 as next-hop must not be considered. This leaves (8 – 1) = 7 routes to nodes and (9 – 3) = 6 routes to networks. 5 out of 7 node destinations, and 4 out of 6 network destinations have an LFA next-hop (L). This leads to (5 + 4) / (7 + 6) = 9/13 IPv4 prefixes that have ISIS IP routing LFA coverage. A similar derivation applies for IPv6 prefixes.

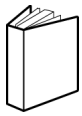
```
[/]
A:admin@PE-4# show router isis routes alternative

===============================================================================
Rtr Base ISIS Instance 0 Route Table (alternative)
===============================================================================
Prefix[Flags]                     Metric     Lvl/Typ     Ver.  SysID/Hostname
  NextHop                                                MT     AdminTag/SID[F]
Alt-Nexthop                                              Alt-   Alt-Type
                                                        Metric
-------------------------------------------------------------------------------
192.0.2.1/32                      10         1/Int.      57    PE-1
   192.168.14.1                                           0       0/1[NnP]
192.0.2.2/32                      20         1/Int.      62    PE-1
   192.168.14.1                                           0       0/2[NnP]
192.0.2.3/32                      20         1/Int.      82    PE-5
   192.168.45.2                                           0       0/3[NnP]
   192.168.46.2(L)                                       70      LP
192.0.2.4/32                      0          1/Int.      3     PE-4
   0.0.0.0                                                0       0/4[NnP]
192.0.2.5/32                      10         1/Int.      82    PE-5
   192.168.45.2                                           0       0/5[NnP]
   192.168.46.2(L)                                       60      LP
192.0.2.6/32                      30         1/Int.      70    PE-6
   192.168.46.2                                           0       0/6[NnP]
   192.168.45.2(L)                                       40      LP
192.0.2.7/32                      30         1/Int.      82    PE-5
   192.168.45.2                                           0       0/7[NnP]
   192.168.46.2(L)                                       40      NP
192.0.2.8/32                      20         1/Int.      82    PE-5
   192.168.45.2                                           0       0/8[NnP]
   192.168.46.2(L)                                       50      NP
192.168.12.0/30                   20         1/Int.      57    PE-1
   192.168.14.1                                           0       0
192.168.14.0/30                   10         1/Int.      6     PE-4
   0.0.0.0                                                0       0
192.168.23.0/30                   40         1/Int.      82    PE-1
   192.168.14.1                                           0       0
192.168.35.0/30                   20         1/Int.      82    PE-5
   192.168.45.2                                           0       0
   192.168.46.2(L)                                       70      LP
192.168.45.0/30                   10         1/Int.      82    PE-4
   0.0.0.0                                                0       0
192.168.46.0/30                   30         1/Int.      66    PE-4
   0.0.0.0                                                0       0
192.168.58.0/30                   20         1/Int.      82    PE-5
   192.168.45.2                                           0       0
   192.168.46.2(L)                                       70      LP
```

```
192.168.67.0/30                  40         1/Int.      82    PE-5
   192.168.45.2                                           0       0
   192.168.46.2(L)                                       50      NP
192.168.78.0/30                  30         1/Int.      82    PE-5
   192.168.45.2                                           0       0
   192.168.46.2(L)                                       60      NP
-------------------------------------------------------------------------------
No. of Routes: 17 (17 paths)
-------------------------------------------------------------------------------
Flags        : L = Loop-Free Alternate nexthop
Alt-Type     : LP = linkProtection, NP = nodeProtection
SID[F]       : R  = Re-advertisement
               N  = Node-SID
               nP = no penultimate hop POP
               E  = Explicit-Null
               V  = Prefix-SID carries a value
               L  = value/index has local significance
===============================================================================
```

Conclusion

TI-LFA extends the calculation of a backup path for cases where the extended P space and the Q space do not overlap. TI-LFA also ensures that the post-failure path coincides with the post-convergence path, which avoids a switchover after SPF convergence.

Customer document and product support



Customer documentation

Customer documentation welcome page

Technical support

Product support portal

Documentation feedback

Customer documentation feedback