



7450 Ethernet Service Switch  
7750 Service Router  
7950 Extensible Routing System  
Releases up to 24.3.R2

## Layer 3 Services Advanced Configuration Guide for Classic CLI

---

3HE 20794 AAAA TQZZA  
Edition: 01  
July 2024

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.

---

This document includes Nokia proprietary and confidential information, which may not be distributed or disclosed to any third parties without the prior written consent of Nokia.

This document is intended for use by Nokia's customers ("You"/"Your") in connection with a product purchased or licensed from any company within Nokia Group of Companies. Use this document as agreed. You agree to notify Nokia of any errors you may find in this document; however, should you elect to use this document for any purpose(s) for which it is not intended, You understand and warrant that any determinations You may make or actions You may take will be based upon Your independent judgment and analysis of the content of this document.

Nokia reserves the right to make changes to this document without notice. At all times, the controlling version is the one available on Nokia's site.

No part of this document may be modified.

NO WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY OF AVAILABILITY, ACCURACY, RELIABILITY, TITLE, NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE, IS MADE IN RELATION TO THE CONTENT OF THIS DOCUMENT. IN NO EVENT WILL NOKIA BE LIABLE FOR ANY DAMAGES, INCLUDING BUT NOT LIMITED TO SPECIAL, DIRECT, INDIRECT, INCIDENTAL OR CONSEQUENTIAL OR ANY LOSSES, SUCH AS BUT NOT LIMITED TO LOSS OF PROFIT, REVENUE, BUSINESS INTERRUPTION, BUSINESS OPPORTUNITY OR DATA THAT MAY ARISE FROM THE USE OF THIS DOCUMENT OR THE INFORMATION IN IT, EVEN IN THE CASE OF ERRORS IN OR OMISSIONS FROM THIS DOCUMENT OR ITS CONTENT.

Copyright and trademark: Nokia is a registered trademark of Nokia Corporation. Other product names mentioned in this document may be trademarks of their respective owners.

© 2024 Nokia.

---

# Table of contents

List of tables.....	5
List of figures.....	6
Preface.....	11
BGP Best External in a VPRN.....	12
Carrier Supporting Carrier IP VPNs.....	35
Flexible Algorithms for SRv6-based VPRNs.....	55
Inter-AS VPRN Model B.....	88
Inter-AS VPRN Model B Using MPLS over UDP.....	106
Inter-AS VPRN Model C.....	119
Intra-AS NG-MVPN over BIER.....	134
Layer 3 VPN: VPRN Type Spoke.....	156
NG-MVPN Configuration with MPLS.....	170
NG-MVPN Configuration with PIM.....	220
NG-MVPN Inter-AS Model B Using Non-Segmented mLDP Tunnels.....	262
NG-MVPN Inter-AS Model C Using Non-Segmented mLDP Tunnels.....	290
NG-MVPN Sender-Only, Receiver-Only.....	322
NG-MVPN Source Redundancy.....	369
NG-MVPN Wildcard S-PMSI.....	398

<b>Rosen MVPN Core Diversity.....</b>	<b>423</b>
<b>Rosen MVPN Inter-AS Option B.....</b>	<b>447</b>
<b>Selective VPRN uRPF Control on Network Interfaces.....</b>	<b>470</b>
<b>Spoke Termination for IPv6-6VPE.....</b>	<b>485</b>
<b>Traffic Leaking from VPRN to GRT.....</b>	<b>506</b>
<b>Weighted ECMP for VPRN over RSVP-TE or SR-TE LSPs.....</b>	<b>525</b>

## List of tables

Table 1: NG MVPN Components.....	221
Table 2: mLDP Message Opaque Value Types in MVPN Model B.....	264
Table 3: mLDP Message Opaque Value Types in MVPN inter-AS Model C.....	292
Table 4: S-PMSI Auto-Discovery BGP NLRI.....	400

## List of figures

Figure 1: CE-4 advertises prefix 10.0.0.0/8 to its EBGP peers PE-1 and PE-2.....	13
Figure 2: Default BGP behavior: BGP advertises best route only.....	14
Figure 3: VPRN BGP best external enabled: BGP advertises active and standby routes.....	15
Figure 4: BGP FRR on PE-1 after failure of active link to CE.....	15
Figure 5: PE-2 re-advertises VPN-IPv4 route with label based on VRF.....	16
Figure 6: Traffic destined for prefix 10.0.0.0/8 after control plane convergence.....	17
Figure 7: Example topology.....	18
Figure 8: Loadsharing for traffic from PE-3 destined to 10.0.0.0/8.....	33
Figure 9: CSC network topology.....	36
Figure 10: Example topology.....	56
Figure 11: Inter-AS VPRN Model B control and data plane example.....	89
Figure 12: Inter-AS VPRN Model B topology.....	90
Figure 13: IP over MPLS over UDP packet format.....	106
Figure 14: Inter-AS VPRN model B topology using MPLS over UDP in AS 64496.....	107
Figure 15: Inter-AS VPN Model C.....	120
Figure 16: Protocol overview.....	121
Figure 17: Bit Forwarding Router types.....	135
Figure 18: BIER control plane: example bit position assignment and advertisement.....	136
Figure 19: BIER data plane: example BIER forwarding table for P-2.....	136
Figure 20: BIER data plane: example BIER packet forwarding.....	137
Figure 21: BIER sets.....	138

---

Figure 22: MVPN over BIER.....	139
Figure 23: PTA: PMSI tunnel attribute.....	139
Figure 24: Intra-AS NG-MVPN over BIER.....	140
Figure 25: CE hub and spoke data path.....	157
Figure 26: CE hub and spoke control plane isolation.....	158
Figure 27: Internal VPRN logic on a PE router.....	158
Figure 28: CE hub and spoke topology and addressing scheme.....	160
Figure 29: Network Topology.....	172
Figure 30: VPRN 1 Topology used for mLDP.....	176
Figure 31: VPRN 2 Topology used for RSVP-TE P2MP.....	194
Figure 32: VPRN 2 Topology used for MVPN Source Redundancy.....	212
Figure 33: VPRN 3 Topology used for MDT SAFI.....	217
Figure 34: Network Topology.....	222
Figure 35: Example Topology for Anycast RP.....	242
Figure 36: IGMP and PIM Control Messaging Schematic.....	247
Figure 37: PIM SSM in Customer Signaling Plane.....	251
Figure 38: NG-MVPN Inter-AS Model B.....	263
Figure 39: Inter-AS MVPN Protocol Requirements.....	265
Figure 40: AS 64496 Protocols.....	265
Figure 41: AS 64497 Protocols.....	268
Figure 42: Inter-AS Protocols.....	270
Figure 43: BGP MVPN Intra-AD Route Advertisement.....	278
Figure 44: P2MP LDP Label Mapping.....	280

---

Figure 45: NG-MVPN Inter-AS Model C.....	291
Figure 46: Inter-AS MVPN Protocol Requirements.....	293
Figure 47: AS 64496 Protocols.....	293
Figure 48: AS 64497 Protocols.....	298
Figure 49: Inter-AS Protocols.....	300
Figure 50: BGP MVPN Intra-AD Route Advertisement.....	310
Figure 51: P2MP LDP Label Mapping.....	312
Figure 52: Default PMSI Hierarchy.....	322
Figure 53: Optimized PMSI Structure.....	323
Figure 54: Example Topology.....	324
Figure 55: RSVP-Based BGP Message Flow Between PE-1 and PE-2.....	332
Figure 56: RSVP-Based BGP Message Flow Between PE-1 and PE-3.....	334
Figure 57: mLDP-Based BGP Message Flow Between PE-1 and PE-2.....	350
Figure 58: mLDP-Based BGP Message Flow Between PE-1 and PE-3.....	356
Figure 59: Source Redundancy Example.....	370
Figure 60: Schematic Topology.....	372
Figure 61: Multicast VPN.....	399
Figure 62: Schematic Topology.....	401
Figure 63: S-PMSI P2MP LSP Schematic.....	418
Figure 64: Core Diversity Schematic.....	425
Figure 65: Core Diversity Network.....	426
Figure 66: Core Diversity Network — Base OSPF.....	427
Figure 67: Core Diversity Network - OSPF Instance 1.....	428



---

Figure 68: General Topology for Inter-AS MVPN.....	447
Figure 69: Protocols Used for Inter-AS MVPN.....	448
Figure 70: BGP Signaling Steps.....	450
Figure 71: PIM-P Signaling Steps for Default MDT.....	450
Figure 72: PIM-C Signaling.....	451
Figure 73: PIM-P Signaling Steps for Data MDT.....	452
Figure 74: Example Topology Details.....	453
Figure 75: BGP Signaling Steps.....	457
Figure 76: PIM-P Signaling Steps for Default MDT.....	460
Figure 77: PIM-C Signaling.....	464
Figure 78: PIM-P Signaling Steps for Data MDT.....	466
Figure 79: Example Topology in AS 64496.....	472
Figure 80: uRPF Enabled in Strict Mode on Access Interface in VPRN 1.....	475
Figure 81: uRPF Checking in Strict Mode in Base Router on PE-1.....	476
Figure 82: uRPF Checking in VPRN 1 on PE-3.....	479
Figure 83: Selective VPRN uRPF on Network Interfaces Enabled for VPRN 1 and Disabled for VPRN 2..	483
Figure 84: Spoke termination for IPv6.....	486
Figure 85: IPv6 addressing and IPv6 prefixes.....	486
Figure 86: MP-BGP VPN IPv6.....	487
Figure 87: Spoke termination for IPv6 addressing.....	489
Figure 88: PE-4 VPRN with SAP to CE-5.....	494
Figure 89: VPRN to GRT leak process.....	507
Figure 90: Example topology with IPv4 addresses.....	507

Figure 91: IPv4 VPRN to GRT route leaking for IS-IS.....	508
Figure 92: Example topology with IPv6 addresses.....	516
Figure 93: Regular ECMP in AS 64496.....	526
Figure 94: Weighted ECMP in AS 64496.....	526
Figure 95: Example Topology.....	528
Figure 96: Weighted ECMP over RSVP LSPs used in a spoke SDP.....	538
Figure 97: Weighted ECMP over SR-TE LSPs in AS 64496.....	542

# 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 24.7.R2. 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 Best External in a VPRN

This chapter provides information about BGP Best External in a VPRN.

Topics in this chapter include:

- [Applicability](#)
- [Overview](#)
- [Configuration](#)
- [Conclusion](#)

## Applicability

The information and configuration in this chapter was originally written for SR OS Release 14.0.R7. In the current edition, the CLI is updated to SR OS Release 22.2.R2.

## Overview

By default, BGP speakers only advertise their best route for a destination. The BGP best external feature allows BGP speakers to advertise their best external route for a prefix/Network Layer Reachability Information (NLRI) to their IBGP peers when their best overall route for this prefix/NLRI is an internal route. This feature provides additional path visibility to the IBGP mesh. When two paths are available to reach a destination, and one is preferred, the availability of an alternate path in the RIB means that only a FIB update is required if the preferred next-hop fails. Also, the presence of two paths can reduce route oscillation.

BGP best external can be enabled in the base router with the following command:

```
*A:PE-2>config>router# bgp ?
- bgp
- no bgp

[no] add-paths      + Enable/Disable BGP ADD-PATHS
[no] advertise-exte* - Enable/Disable Advertise Best External for the bgp family
[no] advertise-inac* - Enable/disable advertising of inactive BGP routes to other BGP
peers
---snip---
```

```
# on PE-2:
configure
  router Base
    bgp
      advertise-external ipv4
```

Chapter *BGP Add-Path* in the *7450 ESS, 7750 SR, and 7950 XRS Unicast Routing Protocols Advanced Configuration Guide for Classic CLI* describes the use of the `add-paths` parameter for different address families. Chapter *BGP Fast Reroute* in the *7450 ESS, 7750 SR, and 7950 XRS Unicast Routing Protocols*

*Advanced Configuration Guide for Classic CLI* includes a configuration example with BGP best external enabled in the base router, whereas this chapter focuses on BGP best external in a **vprn** context.

VPRN BGP best external can be configured with the following command:

```
*A:PE-2>config>service# vprn "VPRN 1" ?
- vprn <service-id> [name <name>] [customer <customer-id>] [create]
- no vprn <service-id>
---snip---
```

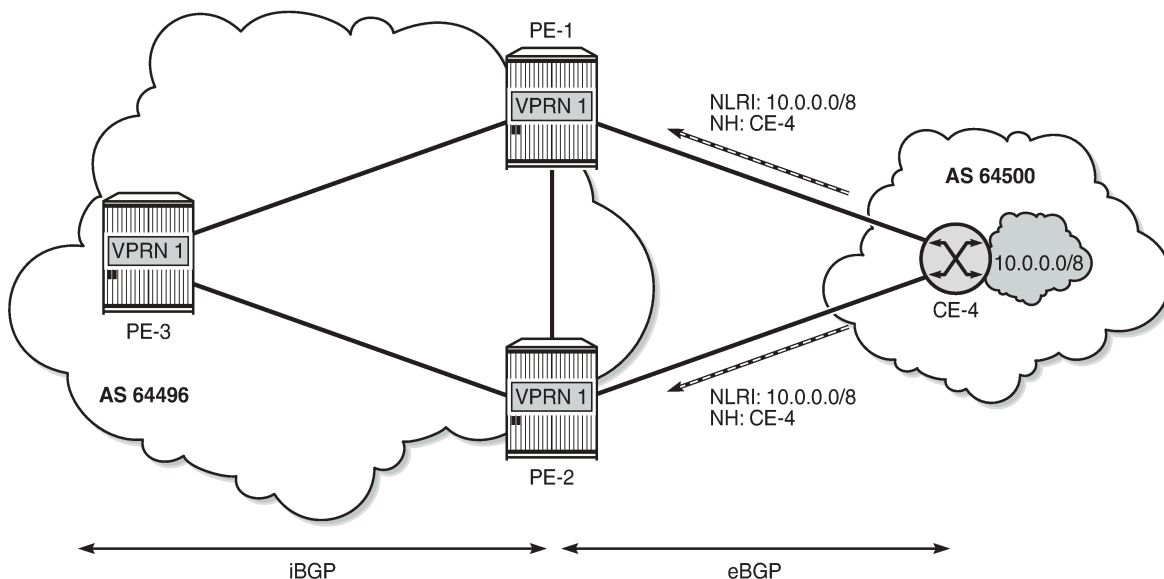
```
[no] export-inactiv* - Allow/Disallow exporting inactive BGP routes
---snip---
```

```
# on PE-2:
configure
service
vprn 1 name "VPRN 1" customer "1" create
export-inactive-bgp
```

VPRN BGP best external allows the best EBGP IPv4/IPv6 route learned by a VPRN to be exported as a BGP VPN-IPv4/IPv6 route, even when that EBGP IPv4/IPv6 route is inactive due to the presence of a preferred BGP VPN-IPv4/IPv6 route from another PE. This best external route advertisement is useful in active/standby multi-homing scenarios because it can ensure that all PEs have knowledge of the backup path provided by the standby PE, thus reducing convergence times. VPRN BGP best external can also be applied in combination with Equal Cost Multi-Path (ECMP).

**Figure 1:** CE-4 advertises prefix 10.0.0.0/8 to its EBGP peers PE-1 and PE-2 shows the example topology with CE-4 in Autonomous System (AS) 64500 advertising prefix 10.0.0.0/8 to VPRN 1 in PE-1 and PE-2 in AS 64496.

*Figure 1: CE-4 advertises prefix 10.0.0.0/8 to its EBGP peers PE-1 and PE-2*

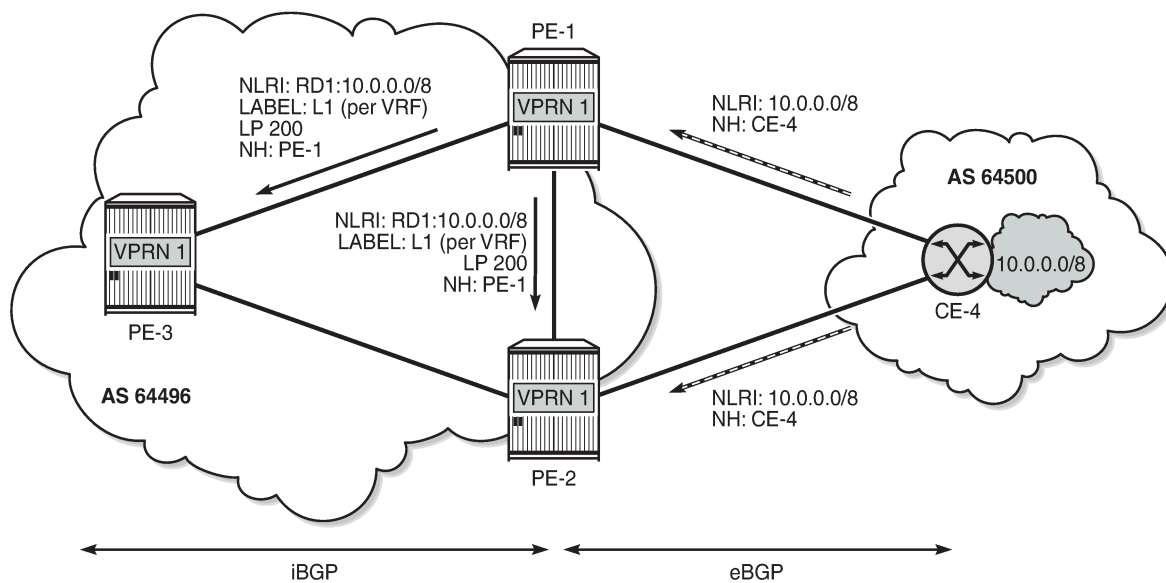


26261

PE-1 is the primary PE for this prefix and it creates a corresponding BGP VPN-IPv4 route with a higher local preference (LP) value (for example, 200) compared to the default LP (100). PE-1 advertises this

BGP VPN-IPv4 route to its IBGP peers PE-2 and PE-3. PE-2 imports this BGP VPN-IPv4 route into its VRF, which deactivates the EBGP route received from CE-4, because it has the default LP of 100 (by BGP selection rules, the highest LP wins). By default, BGP prevents PE-2 from exporting its inactive BGP IPv4 route from CE-4 and, therefore, PE-1 and PE-3 cannot learn a BGP VPN-IPv4 backup route for prefix 10.0.0.0/8, as shown in [Figure 2: Default BGP behavior: BGP advertises best route only](#).

Figure 2: Default BGP behavior: BGP advertises best route only



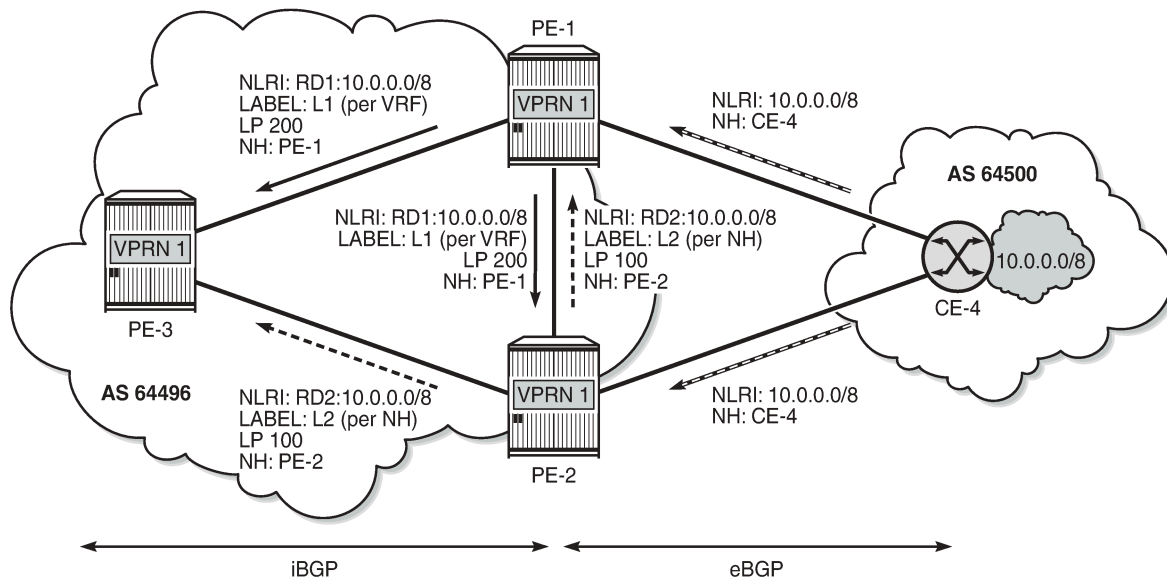
26262

VPRN BGP best external allows PE-2 to advertise its best external route as backup on the following conditions:

- The option **export-inactive-bgp** is configured in VPRN 1 on PE-2 (or on all PEs in the multi-homed site).
- The BGP route from CE-4 must match the VRF export policy in PE-2.
- The BGP VPN-IPv4 route exported by PE-2 must have a unique NLRI (RD:IP prefix combination) that does not overlap with a BGP VPN-IPv4 route from another PE for the same prefix. Therefore, a different RD can be allocated to the VRF in each PE connected to the multi-homed site. For example, VPRN 1 in PE-1 has RD 64496:11 and VPRN 1 in PE-2 has RD 64496:12.

[Figure 3: VPRN BGP best external enabled: BGP advertises active and standby routes](#) shows the BGP route advertisements when VPRN BGP best external is enabled. The BGP VPN-IPv4 route from PE-2 carries a per-next-hop label (meaning pop and forward to CE-4) regardless of the configured label mode of the VPRN service in PE-2.

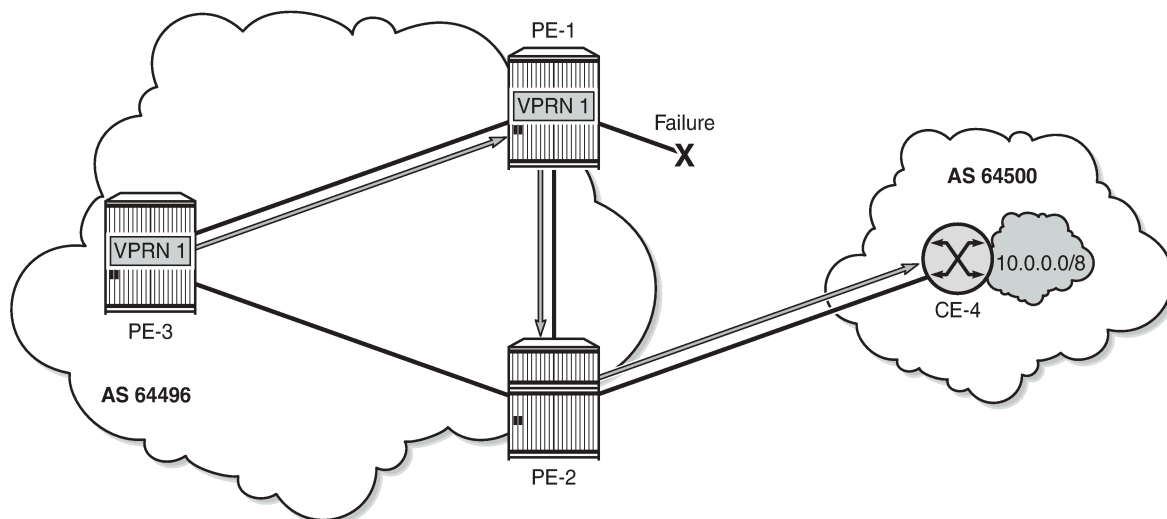
Figure 3: VPRN BGP best external enabled: BGP advertises active and standby routes



26263

The PEs support BGP Fast Reroute (BGP FRR) using BGP VPN-IPv4 routes; therefore, PE-1 and PE-3 can install the route advertised by PE-2 as a backup path for prefix 10.0.0.0/8 and use that path immediately after detecting that the primary path has failed. When the link between PE-1 and CE-4 fails, PE-1 will detect this link failure typically seconds before the other PEs do. Therefore, PE-3 keeps sending traffic toward the network 10.0.0.0/8 to PE-1 and PE-1 uses the repair path via PE-2, as shown in [Figure 4: BGP FRR on PE-1 after failure of active link to CE](#).

Figure 4: BGP FRR on PE-1 after failure of active link to CE

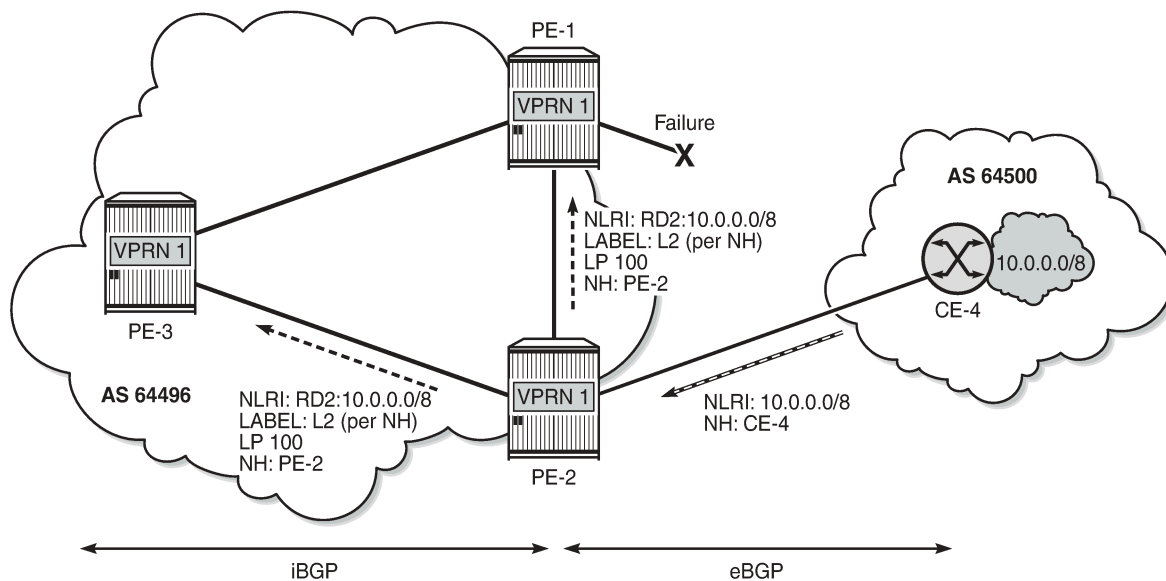


26264

Even when PE-2 is still unaware of the link failure between PE-1 and CE-4, PE-2 will not loop traffic back to PE-1. The reason is that PE-1 sends traffic to PE-2 with a per-next-hop label so that no FIB lookup occurs in PE-2. Traffic is forwarded correctly to CE-4.

When PE-2 receives the BGP VPN-IPv4 route withdrawal from PE-1 for prefix 10.0.0.0/8, it removes the route from its RIB-IN and reruns the decision process. In this example, the EBGP route to CE-4 becomes the new primary/best path. PE-2 will re-advertise its BGP VPN-IPv4 route for prefix 10.0.0.0/8. The difference is that the BGP VPN-IPv4 route is based on the export of an active/used route and, therefore, the advertised label value is based on the configured label mode of the VPRN service, as shown in [Figure 5: PE-2 re-advertises VPN-IPv4 route with label based on VRF](#) for label mode VRF (default).

Figure 5: PE-2 re-advertises VPN-IPv4 route with label based on VRF



26265

It takes time for this route to reach all ingress routers and for these routers to update their forwarding tables to use the per-VRF label value. For a while, there may still be traffic destined for prefix 10.0.0.0/8 that is received by PE-2 with the per-next-hop label L2. Traffic will be dropped if the per-next-hop label is deleted by the IOM as soon as PE-2 determines there are no more inactive/standby paths with CE-4 as next hop. Traffic loss can be avoided by delaying the deletion of per-next-hop labels in the IOM by configuring label retention for BGP labels with the following command:

```
*A:PE-2>config>router>mpls-labels# bgp-labels-hold-timer ?
- bgp-labels-hold-timer <seconds>
- no bgp-labels-hold-timer

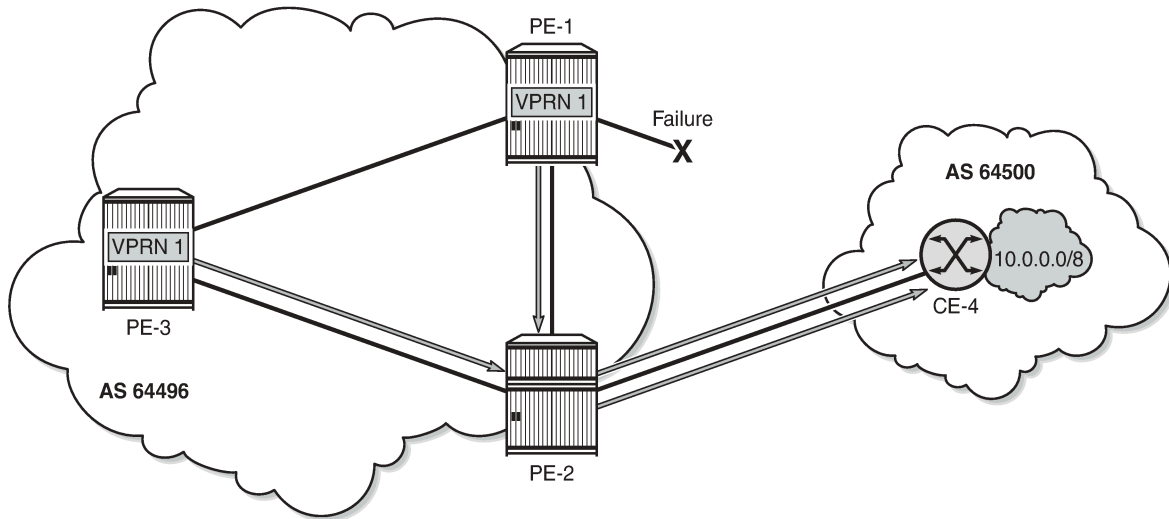
<seconds>          : [0..255]
```

```
# on PE-2:
configure
router Base
 mpls-labels
  bgp-labels-hold-timer 60
```



Finally, all ingress routers have updated their forwarding tables based on the BGP update sent by PE-2, and PE-3 sends traffic for prefix 10.0.0.0/8 directly toward PE-2, as shown in [Figure 6: Traffic destined for prefix 10.0.0.0/8 after control plane convergence](#).

Figure 6: Traffic destined for prefix 10.0.0.0/8 after control plane convergence

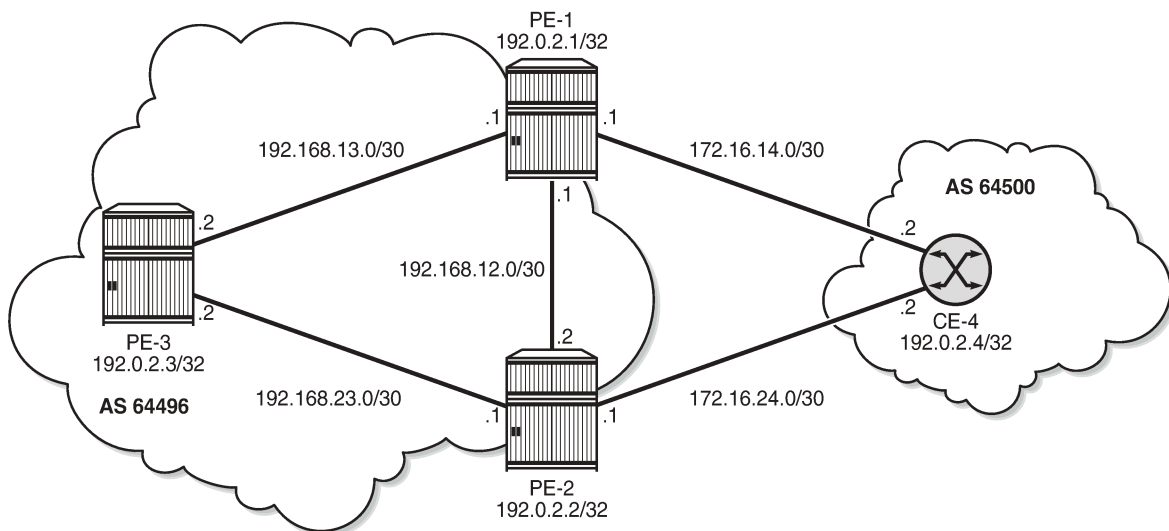


26266

## Configuration

[Figure 7: Example topology](#) shows the example topology with the used IP addresses.

Figure 7: Example topology



26267

The initial configuration includes the following:

- Cards, MDAs, ports
- Router interfaces
- IS-IS (or OSPF) as IGP within AS 64496
- LDP on all interfaces within AS 64496

BGP is configured in the base router context of all PEs for address family VPN-IPv4; for example, for PE-1 as follows:

```
# on PE-1:
configure
router Base
  autonomous-system 64496
  bgp
    rapid-withdrawal
    group "IBGP"
      family vpn-ipv4
      peer-as 64496
      neighbor 192.0.2.2
    exit
    neighbor 192.0.2.3
    exit
  exit
```

The BGP configuration for the base router on the other two PEs is similar and a full mesh is established in AS 64496.

## Configure VPRN without BGP best external

VPRN 1 is created on all PEs with the following settings:

- Default label mode: label-mode vrf
- Ready for BGP FRR: **enable-bgp-vpn-backup ipv4**
- Different RDs in VPRN 1 for each PE: 64496:11 on PE-1, 64496:12 on PE-2, and 64496:13 on PE-3
- Auto-bind-tunnel with resolution any. In this example, LDP will be used.
- Loopback interface "lo0" with IP address 172.31.2.1/32 on PE-1, which is also defined as the router ID in VPRN 1. The same approach is used on PE-2 and PE-3: 172.31.2.2/32 and 172.31.2.3/32.
- IBGP between all PEs (full mesh) for address family IPv4
- EBGP between PE-1 and CE-4 and between PE-2 and CE-4
- BGP best external is disabled, by default.

The configuration of VPRN 1 on PE-3 is as follows:

```
# on PE-3:
configure
  service
    vprn 1 name "VPRN 1" customer 1 create
      router-id 172.31.2.3
      autonomous-system 64496
      label-mode vrf # default
      enable-bgp-vpn-backup ipv4
      interface "lo0" create
        address 172.31.2.3/32
        loopback
      exit
      bgp-ipvpn
        mpls
          auto-bind-tunnel
            resolution any
          exit
          route-distinguisher 64496:13
          vrf-target target:64496:1
          no shutdown
        exit
      exit
    bgp
      rapid-withdrawal
      group "IBGP"
        peer-as 64496
        neighbor 172.31.2.1
        exit
        neighbor 172.31.2.2
        exit
      exit
    exit
  no shutdown
```

On PE-1 and PE-2, the VPRN configuration includes an external interface toward CE-4, and EBGP is defined toward peer CE-4. The VPRN 1 configuration on PE-2 is as follows:

```
# on PE-2:
configure
  service
    vprn 1 name "VPRN 1" customer 1 create
      router-id 172.31.2.2
      autonomous-system 64496
      label-mode vrf # default
      enable-bgp-vpn-backup ipv4
```

```

interface "lo0" create
  address 172.31.2.2/32
  loopback
exit
interface "int-PE-2-CE-4_VPRN1" create
  address 172.16.24.1/30
  sap 1/1/3:1 create
  exit
exit
bgp-ipvpn
  mpls
    auto-bind-tunnel
    resolution any
  exit
  route-distinguisher 64496:12
  vrf-target target:64496:1
  no shutdown
  exit
exit
bgp
  rapid-withdrawal
  split-horizon
  group "EBGP"
    peer-as 64500
    neighbor 172.16.24.2
  exit
  group "IBGP"
    peer-as 64496
    neighbor 172.31.2.1
  exit
  neighbor 172.31.2.3
  exit
  exit
  no shutdown

```

PE-2 does not have an import policy that sets the LP and, therefore, the default LP of 100 is used for routes imported from EBGP peer CE-4.

The VPRN 1 configuration on PE-1 looks similar to the configuration on PE-2, but includes an import policy that assigns an LP of 200 to each prefix that is received from CE-4, as follows:

```

# on PE-1:
configure
  router Base
    policy-options
      begin
        policy-statement "import-bgp-LP200"
          default-action accept
          local-preference 200
        exit
      exit
    exit
  commit
  exit
  service
    vprn 1 name "VPRN 1" customer 1 create
    router-id 172.31.2.1
    autonomous-system 64496
    label-mode vrf # default
    enable-bgp-vpn-backup ipv4
    interface "lo0" create

```

```

        address 172.31.2.1/32
        loopback
    exit
    interface "int-PE-1-CE-4_VPRN1" create
        address 172.16.14.1/30
        sap 1/1/3:1 create
        exit
    exit
    bgp-ipvpn
        mpls
            auto-bind-tunnel
                resolution any
            exit
            route-distinguisher 64496:11
            vrf-target target:64496:1
            no shutdown
        exit
    exit
    bgp
        rapid-withdrawal
        split-horizon
        group "EBGP"
            import "import-bgp-LP200"
            peer-as 64500
            neighbor 172.16.14.2
            exit
        exit
        group "IBGP"
            peer-as 64496
            neighbor 172.31.2.2
            exit
            neighbor 172.31.2.3
            exit
        exit
    exit
    no shutdown

```

CE-4 has EBGP configured toward PE-1 and PE-2. CE-4 exports the prefix 10.0.0.0/8, as defined in export policy "export-bgp" that is applied in the **bgp** context:

```

# on CE-4:
configure
    router Base
        interface "int-CE-4-PE-1_VPRN1"
            address 172.16.14.2/30
            port 1/1/1:1
        exit
        interface "int-CE-4-PE-2_VPRN1"
            address 172.16.24.2/30
            port 1/1/2:1
        exit
        interface "system"
            address 192.0.2.4/32
        exit
        interface "test_connectedNW"
            address 10.0.0.1/8
            loopback
        exit
        autonomous-system 64500
        policy-options
            begin
                prefix-list "10.0.0.0/8"
                prefix 10.0.0.0/8 longer
            end
    end

```

```

    exit
    policy-statement "export-bgp"
    entry 10
    from
    prefix-list "10.0.0.0/8"
    exit
    action accept
    exit
    exit
    exit
    commit
exit
bgp
    rapid-withdrawal
    split-horizon
    group "EBGP"
    export "export-bgp"
    peer-as 64496
    neighbor 172.16.14.1
    exit
    neighbor 172.16.24.1
    exit
    exit
exit
exit

```

Initially, VPRN BGP best external is disabled and, so only the best BGP route will be advertised and IBGP peers will not learn backup paths. The following section shows which routes are exchanged. Afterward, VPRN BGP best external will be enabled and the same show commands will be used.

## Verification - VPRN without BGP best external

PE-1 imports prefix 10.0.0.0/8, assigns LP 200 to it, and advertises a corresponding VPN-IPv4 route to its IBGP peers (PE-2 and PE-3). Toward PE-2, this is as follows:

```

# on PE-1:
9 2022/04/29 09:56:15.585 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 = 65
  Flag: 0x90 Type: 14 Len: 30 Multiprotocol Reachable NLRI:
    Address Family VPN_IPV4
    NextHop len 12 NextHop 192.0.2.1
    10.0.0.0/8 RD 64496:11 Label 524284 (Raw label 0x7ffc1)
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 6 AS Path:
    Type: 2 Len: 1 < 64500 >
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 200
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:64496:1
"

```

The NLRI includes the prefix 10.0.0.0/8 and the RD 64496:11, and the label is 524284. BGP prevents PE-2 from sending a similar BGP update for prefix 10.0.0.0/8 because that route is not active on PE-2. PE-3 receives a BGP VPN-IPv4 route for network 64496:11:10.0.0.0/8, and this route has PE-1 as next hop and LP 200. No route is received from PE-2 for network 64496:12:10.0.0.0/8; as follows:

```

*A:PE-3# show router bgp routes vpn-ipv4
=====

```

```

BGP Router ID:192.0.2.3      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
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 64496:11:10.0.0.0/8                    200      None
      192.0.2.1                             None      10
      64500                                  524284
u*>i 64496:11:172.16.14.0/30                 100      None
      192.0.2.1                             None      10
      No As-Path                             524284
u*>i 64496:11:172.31.2.1/32                 100      None
      192.0.2.1                             None      10
      No As-Path                             524284
u*>i 64496:12:172.16.24.0/30                100      None
      192.0.2.2                             None      10
      No As-Path                             524284
u*>i 64496:12:172.31.2.2/32                 100      None
      192.0.2.2                             None      10
      No As-Path                             524284
-----
Routes : 5
=====

```

In a similar way, the list of BGP VPN-IPv4 routes on PE-2 includes prefix 64496:11:10.0.0.0/8 with LP 200 and next hop PE-1, as follows:

```

*A:PE-2# show router bgp routes vpn-ipv4
=====
BGP Router ID:192.0.2.2      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
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 64496:11:10.0.0.0/8                    200      None
      192.0.2.1                             None      10
      64500                                  524284
u*>i 64496:11:172.16.14.0/30                 100      None
      192.0.2.1                             None      10
      No As-Path                             524284
u*>i 64496:11:172.31.2.1/32                 100      None
      192.0.2.1                             None      10
      No As-Path                             524284
u*>i 64496:13:172.31.2.3/32                 100      None
      192.0.2.3                             None      10

```

```

No As-Path                                     524284
-----
Routes : 4
=====

```

The list of BGP IPv4 routes in VPRN 1 on PE-2 has two entries for prefix 10.0.0.0/8, but none of them is best or used, as follows:

```

*A:PE-2# show router 1 bgp routes
=====
BGP Router ID:172.31.2.2      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP IPv4 Routes
=====
Flag  Network                               LocalPref  MED
     Nexthop (Router)                       Path-Id    IGP Cost
     As-Path                                Label
-----
*i   10.0.0.0/8                               None       None
     172.16.24.2                               None       0
     64500                                       -
i    10.0.0.0/8                               200       None
     172.16.14.2                               None       0
     64500                                       -
-----
Routes : 2
=====

```

The routing table for VPRN 1 on PE-2 and PE-3 for prefix 10.0.0.0/8 shows that the next hop is PE-1 and the protocol is BGP VPN, as follows:

```

*A:PE-2# show router 1 route-table 10.0.0.0/8
=====
Route Table (Service: 1)
=====
Dest Prefix[Flags]                Type   Proto   Age           Pref
  Next Hop[Interface Name]                Metric
-----
10.0.0.0/8                          Remote BGP VPN 00h01m53s 170
  192.0.2.1 (tunneled)                  10
-----
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
=====

```

Instead of using an external route to CE-4, the route for prefix 10.0.0.0/8 is internal (BGP VPN), using an LDP transport tunnel to PE-1. There are no non-active routes, as can be shown by adding the keyword **all** to the preceding show command, as follows:

```

*A:PE-2# show router 1 route-table 10.0.0.0/8 all

```



```

=====
Route Table (Service: 1)
=====
Dest Prefix[Flags]                Type   Proto   Age      Pref
  Next Hop[Interface Name]         Active Metric
-----
10.0.0.0/8                          Remote BGP VPN 00h02m11s 170
  192.0.2.1 (tunneled)              Y      10
-----
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
       E = Inactive best-external BGP route
=====

```

There are no standby routes, because BGP only advertises the best used route.

On PE-1, the following BGP IPv4 route with next hop CE-4 is used for prefix 10.0.0.0/8 in VPRN 1:

```

*A:PE-1# show router 1 bgp routes
=====
BGP Router ID:172.31.2.1      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP IPv4 Routes
=====
Flag  Network                LocalPref  MED
      Nexthop (Router)      Path-Id    IGP Cost
      As-Path                Label
-----
u*>i 10.0.0.0/8                200        None
      172.16.14.2           None        0
      64500                  -
-----
Routes : 1
=====

```

The route for prefix 10.0.0.0/8 in the routing table of VPRN 1 has next hop 172.16.14.2 on CE-4, as follows:

```

*A:PE-1# show router 1 route-table 10.0.0.0/8 all
=====
Route Table (Service: 1)
=====
Dest Prefix[Flags]                Type   Proto   Age      Pref
  Next Hop[Interface Name]         Active Metric
-----
10.0.0.0/8                          Remote BGP     00h03m06s 170
  172.16.14.2                          Y      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
=====

```

E = Inactive best-external BGP route

There is no backup route because BGP prevents PE-2 from sending a standby route for prefix 10.0.0.0/8 to its IBGP peers.

PE-2 has advertised two VPN-IPv4 routes in the base router (the last number in Rcv/Act/Sent = Received/Active/Sent), as follows:

```
*A:PE-2# show router bgp summary family vpn-ipv4
=====
BGP Router ID:192.0.2.2      AS:64496      Local AS:64496
=====
BGP Admin State      : Up      BGP Oper State      : Up
---snip---
=====
BGP VPN-IPv4 Summary
=====
Legend : D - Dynamic Neighbor
=====
Neighbor
      AS PktRcvd PktSent  InQ  OutQ  Up/Down  State|Recv/Actv/Sent
-----
192.0.2.1
      64496      18      17    0    0 00h06m02s 3/3/2
192.0.2.3
      64496      16      17    0    0 00h05m55s 1/1/2
-----
```

## Enable BGP best external in VPRN

VPRN BGP best external is configured on PE-2 (or on all PEs in the multi-homing site) as follows:

```
# on PE-2:
configure
  service
    vprn "VPRN 1"
      export-inactive-bgp
```

When configured, this command causes all IPv4 and IPv6 VPRN BGP best external routes to be exported in the multi-protocol BGP (MP-BGP) domain. Best external routes are BGP routes for which all the following conditions are met:

- The BGP route is matched by the VRF export policy.
- The BGP route is inactive because a more preferred BGP VPN route for the same prefix is present in the route table manager (RTM).
- This BGP route is best and valid considering only VPRN BGP routes.

PE-2 is advertising a best external route and is called the standby PE for prefix 10.0.0.0/8. PEs can be active for some IP prefixes and standby for other IP prefixes.

Best external routes are advertised to the BGP VPN-IPv4 neighbors. In this example, the BGP VPN-IPv4 neighbors are IBGP neighbors, but they can also be EBGP neighbors. The RD must be unique across the PEs exporting a BGP VPN-IP route for the same prefix; otherwise, the best external route may not be advertised. The advertised VPRN label is based on the next hop IP of the best external route, regardless of the label mode of the VPRN in the standby PE.

## Verification - VPRN with BGP best external - BGP FRR

VPRN with BGP best external BGP FRR results in the following. VPRN BGP best external is enabled (BGP Export Inactv) in VPRN 1 on PE-2:

```
*A:PE-2# show service id 1 base

=====
Service Basic Information
=====
Service Id       : 1                Vpn Id           : 0
Service Type    : VPRN
MACSec enabled  : no
Name            : VPRN 1
Description     : (Not Specified)
Customer Id     : 1                Creation Origin  : manual
---snip---

Max IPv6 Routes : No Limit
Ignore NH Metric : Disabled
Hash Label      : Disabled
Entropy Label   : Disabled
Vrf Target      : target:64496:1
---snip---

Label mode      : vrf
BGP VPN Backup  : ipv4
BGP Export Inactv : Enabled
LOG all events  : Disabled

SAP Count       : 1                SDP Bind Count   : 0
VSD Domain     : <none>

-----
Service Access & Destination Points
-----
Identifier                               Type      AdmMTU  OprMTU  Adm  Opr
-----
sap:1/1/3:1                              q-tag    1578   1578   Up   Up
=====
```

After VPRN BGP best external is enabled, PE-2 advertises its standby route for prefix 10.0.0.0/8 to its IBGP peers, as follows:

```
# on PE-2:
16 2022/04/29 10:00:35.266 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.3
"Peer 1: 192.0.2.3: UPDATE
Peer 1: 192.0.2.3 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 65
  Flag: 0x90 Type: 14 Len: 30 Multiprotocol Reachable NLRI:
    Address Family VPN_IPV4
    NextHop len 12 NextHop 192.0.2.2
    10.0.0.0/8 RD 64496:12 Label 524283 (Raw label 0x7fffb1)
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 6 AS Path:
    Type: 2 Len: 1 < 64500 >
Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:64496:1
```

The RD is 64496:12, the LP is 100, and the label is 524283. The BGP update shown is sent by PE-2 toward PE-3; the BGP update sent by PE-2 toward PE-1 is similar.

The number of BGP VPN-IPv4 routes sent by PE-2 to each IBGP peer increased from 2 to 3, as follows:

```
*A:PE-2# 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
-----
192.0.2.1
Def. Inst          64496      24   0 00h09m12s 3/3/3 (VpnIPv4)
                24   0
192.0.2.3
Def. Inst          64496      23   0 00h09m06s 1/1/3 (VpnIPv4)
                25   0
---snip---
```

PE-3 has two BGP VPN-IPv4 routes for prefix 10.0.0.0/8: one for network 64496:11:10.0.0.0/8 with LP 200 and next hop PE-1, and one for network 64496:12:10.0.0.0/8 with LP 100 and next hop PE-2, as follows:

```
*A:PE-3# show router bgp routes 10.0.0.0/8 vpn-ipv4
=====
BGP Router ID:192.0.2.3      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
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  64496:11:10.0.0.0/8                     200        None
      192.0.2.1                           None        10
      64500                                 524284
u*>i  64496:12:10.0.0.0/8                     100        None
      192.0.2.2                           None        10
      64500                                 524283
-----
Routes : 2
=====
```

PE-1 has one BGP VPN-IPv4 route for network 64496:12:10.0.0.0/8 with LP 100 and next hop PE-2; PE-2 has one BGP VPN-IPv4 route for network 64496:11:10.0.0.0/8 with LP 200 and next hop PE-1.

All PEs are ready for BGP FRR and the "B" flag indicates that a BGP VPN-IPv4 backup route is available. This flag is present when the VPRN is configured for BGP FRR (**enable-bgp-vpn-backup**) and a standby

route has been received, as follows. The B flag was not present in the output for the routing table when VPRN BGP best external was disabled, as shown earlier.

```
*A:PE-1# show router 1 route-table 10.0.0.0/8

=====
Route Table (Service: 1)
=====
Dest Prefix[Flags]                Type   Proto   Age           Pref
  Next Hop[Interface Name]                Metric
-----
10.0.0.0/8 [B]                    Remote BGP     00h03m17s    170
  172.16.14.2                        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
=====
```

The active route on PE-1 has next hop 172.16.14.2 on CE-4.

On PE-3, the active BGP VPN-IPv4 route for prefix 10.0.0.0/8 uses an LDP transport tunnel to PE-1; a BGP VPN-IPv4 backup route is also available, as follows:

```
*A:PE-3# show router 1 route-table 10.0.0.0/8

=====
Route Table (Service: 1)
=====
Dest Prefix[Flags]                Type   Proto   Age           Pref
  Next Hop[Interface Name]                Metric
-----
10.0.0.0/8 [B]                    Remote BGP VPN  00h06m47s    170
  192.0.2.1 (tunneled)                10
-----
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
=====
```

The active BGP VPN-IPv4 route on PE-2 uses an LDP transport tunnel to PE-1, but no BGP backup route is available:

```
*A:PE-2# show router 1 route-table 10.0.0.0/8

=====
Route Table (Service: 1)
=====
Dest Prefix[Flags]                Type   Proto   Age           Pref
  Next Hop[Interface Name]                Metric
-----
10.0.0.0/8                        Remote BGP VPN  00h07m10s    170
  192.0.2.1 (tunneled)                10
-----
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

PE-2 has a standby BGP IPv4 route that is displayed with the following show command:

```
*A:PE-2# show router 1 route-table 10.0.0.0/8 all
```

```
=====
```

```
Route Table (Service: 1)
```

```
=====
```

Dest Prefix[Flags] Next Hop[Interface Name]	Type	Proto Active	Age Metric	Pref
10.0.0.0/8 [E] 172.16.24.2	Remote	BGP N	00h04m12s 0	170
10.0.0.0/8 192.0.2.1 (tunneled)	Remote	BGP VPN Y	00h08m04s 10	170

```
-----
```

```
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
```

```
E = Inactive best-external BGP route
```

```
=====
```

The "E" flag indicates that this route is an inactive best external BGP route.

VPRN 1 on PE-1 and PE-3 is ready for BGP FRR (**enable-bgp-vpn-backup**) and PE-2 advertised a standby BGP VPN-IPv4 route for prefix 10.0.0.0/8; therefore, PE-1 and PE-3 can add an alternative route to the routing table of VPRN 1, as follows:

```
*A:PE-1# show router 1 route-table 10.0.0.0/8 alternative
```

```
=====
```

```
Route Table (Service: 1)
```

```
=====
```

Dest Prefix[Flags] Next Hop[Interface Name] Alt-NextHop	Type	Proto	Age Metric Alt- Metric	Pref
10.0.0.0/8 172.16.14.2	Remote	BGP	00h04m32s 0	170
<b>10.0.0.0/8 (Backup)</b> <b>192.0.2.2 (tunneled)</b>	<b>Remote</b>	<b>BGP VPN</b>	<b>00h04m32s</b> <b>10</b>	<b>170</b>

```
-----
```

```
No. of Routes: 2
```

```
Flags: n = Number of times nexthop is repeated
```

```
Backup = BGP backup route
```

```
LFA = Loop-Free Alternate nexthop
```

```
S = Sticky ECMP requested
```

```
=====
```

```
*A:PE-3# show router 1 route-table 10.0.0.0/8 alternative
```

```
=====
```

```
Route Table (Service: 1)
```

```
=====
```

Dest Prefix[Flags] Next Hop[Interface Name]	Type	Proto	Age Metric	Pref
--	------	-------	---------------	------

Alt-NextHop		Alt-Metric
10.0.0.0/8	Remote BGP VPN	00h08m04s 170
192.0.2.1 (tunneled)		10
<b>10.0.0.0/8 (Backup)</b>	<b>Remote BGP VPN</b>	<b>00h08m04s 170</b>
<b>192.0.2.2 (tunneled)</b>		<b>10</b>

-----

No. of Routes: 2  
 Flags: n = Number of times nexthop is repeated  
**Backup = BGP backup route**  
 LFA = Loop-Free Alternate nexthop  
 S = Sticky ECMP requested

=====

The alternative BGP VPN-IPv4 route for prefix 10.0.0.0/8 in VPRN 1 uses an LDP transport tunnel toward PE-2.

### Configure ECMP

Because BGP best external allows advertising of an alternative path, it can also be used for load-sharing. ECMP is configured with value 2 in VPRN 1 on all PEs, as follows:

```
# on PE-1, PE-2, PE-3:
configure
  service
    vprn "VPRN 1"
    ecmp 2
```

Other than the ECMP configuration, the VPRN configuration is the same as in the previous example. If ECMP is configured, BGP FRR is not needed anymore:

```
# on PE-1, PE-2, PE-3:
configure
  service
    vprn "VPRN 1"
    no enable-bgp-vpn-backup
```

On PE-3, the BGP decision process will prefer the route with the highest LP and, therefore, only the route via PE-1 with LP 200 will be used and there will be no load-sharing. To ensure that the routes via PE-1 and PE-2 have the same cost, the import policy in VPRN 1 on PE-1 that sets the LP to 200 is removed, as follows:

```
# on PE-1:
configure
  service
    vprn "VPRN 1"
    bgp
      group "EBGP"
      no import
```

BGP best external is enabled (on PE-1 and) PE-2, as follows:

```
# on PE-2:
configure
  service
    vprn "VPRN 1"
```

**export-inactive-bgp**

**Verification - VPRN with BGP best external - ECMP**

VPRN with BGP best external ECMP results in the following. With BGP best external enabled on the PEs in the multi-homing site (PE-2 and PE-3), the following two BGP VPN-IPv4 routes are used on PE-3:

```
*A:PE-3# show router bgp routes 10.0.0.0/8 vpn-ipv4
=====
BGP Router ID:192.0.2.3      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP VPN-IPv4 Routes
=====
Flag  Network                               LocalPref  MED
      Nexthop (Router)                     Path-Id     IGP Cost
      As-Path                               Path-Id     Label
-----
u*>i  64496:11:10.0.0.0/8                    100        None
      192.0.2.1                             None        10
      64500                                   524284
u*>i  64496:12:10.0.0.0/8                    100        None
      192.0.2.2                             None        10
      64500                                   524284
-----
Routes : 2
=====
```

The following BGP IPv4 routes are learned in VPRN 1 on PE-3, but they are not used:

```
*A:PE-3# show router 1 bgp routes 10.0.0.0/8
=====
BGP Router ID:172.31.2.3    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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP IPv4 Routes
=====
Flag  Network                               LocalPref  MED
      Nexthop (Router)                     Path-Id     IGP Cost
      As-Path                               Path-Id     Label
-----
i     10.0.0.0/8                             100        None
      172.16.14.2                           None        0
      64500                                   -
i     10.0.0.0/8                             100        None
      172.16.24.2                           None        0
      64500                                   -
-----
Routes : 2
=====
```



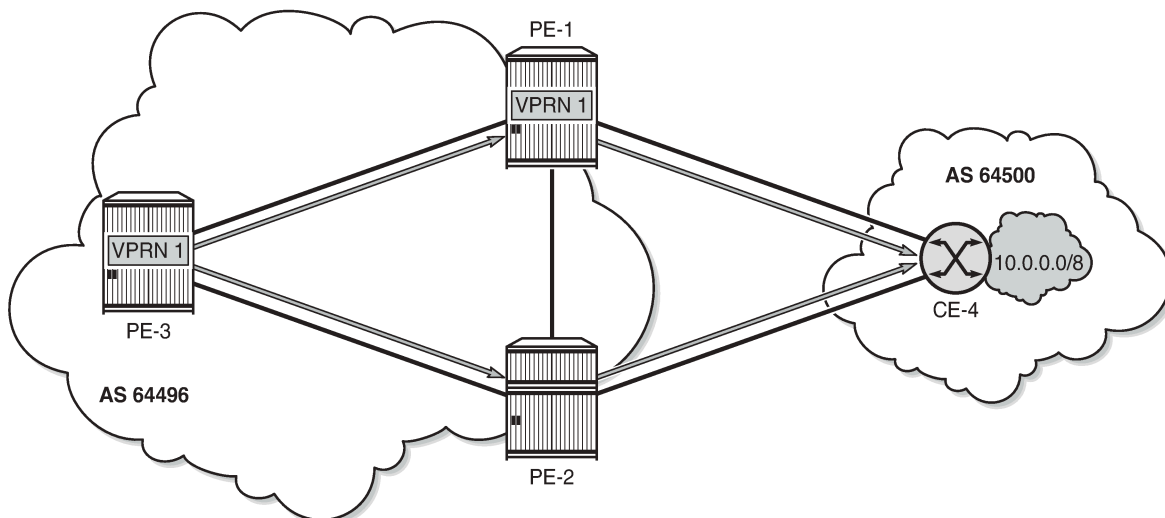
When ECMP is enabled and the routes have the same LP, the routing table on PE-3 has two active routes for prefix 10.0.0.0/8, each using an LDP transport tunnel, as follows:

```
*A:PE-3# show router 1 route-table 10.0.0.0/8

=====
Route Table (Service: 1)
=====
Dest Prefix[Flags]                               Type  Proto  Age      Pref
Next Hop[Interface Name]                       Metric
-----
10.0.0.0/8                                       Remote BGP VPN 00h01m50s 170
192.0.2.1 (tunneled)                             10
10.0.0.0/8                                       Remote BGP VPN 00h01m50s 170
192.0.2.2 (tunneled)                             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
=====
```

Figure 8: Loadsharing for traffic from PE-3 destined to 10.0.0.0/8 shows that traffic from VPRN 1 on PE-3 destined to prefix 10.0.0.0/8 is sprayed over two paths: one via PE-1 and one via PE-2.

Figure 8: Loadsharing for traffic from PE-3 destined to 10.0.0.0/8



26268

## Conclusion

VPRNs can be configured with the option **export-inactive-bgp**, which allows a BGP speaker to advertise its best external BGP route to its BGP peers even if that route is inactive due to the presence of a more preferred BGP VPN route from another PE. BGP best external in VPRN is useful in active/standby multi-homing scenarios because it allows the standby PE to advertise a backup path. The traffic failover time can

be reduced when all PE routers have advance knowledge of the potential backup paths and do not have to wait for BGP route advertisements and/or withdrawals to reprogram their forwarding tables. VPRN BGP best external can also be used in combination with ECMP.

# Carrier Supporting Carrier IP VPNs

This chapter provides information about carrier supporting carrier IP VPN configurations.

Topics in this chapter include:

- [Applicability](#)
- [Overview](#)
- [Configuration](#)
- [Conclusion](#)

## Applicability

This chapter was initially written for on SR OS Release 11.0.R1. The CLI in the current edition corresponds to SR OS Release 22.2.R1. Carrier Supporting Carrier is supported on the 7750 SR and 7950 XRS.

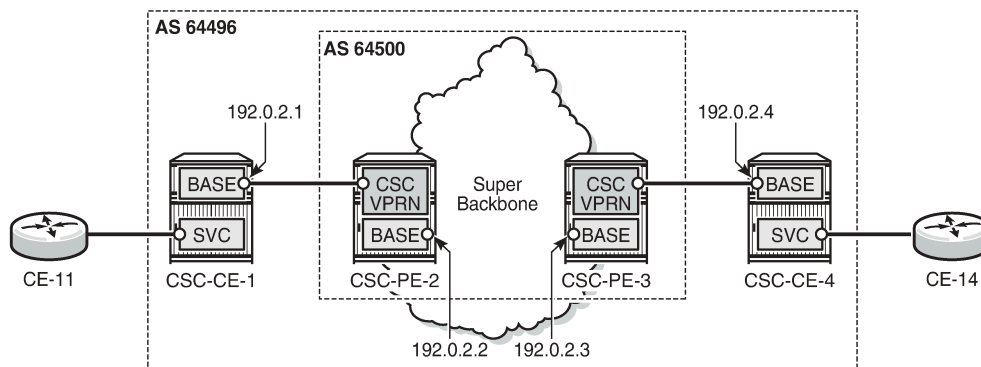
## Overview

Carrier Supporting Carrier (CSC) is a solution that allows one service provider (the Customer Carrier) to use the IP VPN service of another service provider (the Super Carrier) for some or all of its backbone transport. RFC 4364 defines a Carrier Supporting Carrier solution for BGP/MPLS IP VPNs that uses MPLS at the interconnection points between the two service providers to provide a scalable and secure solution.

A simplified CSC network topology is shown in [Figure 9: CSC network topology](#). A CSC deployment involves the following types of devices:

- CE — Customer premises equipment dedicated to one enterprise.
- PE — Edge router managed and operated by the Customer Carrier that connects to CEs to provide business VPN or Internet services.
- CSC-CE — Peering router managed and operated by the Customer Carrier that is connected to CSC-PEs for purposes of using the associated CSC IP VPN services for backbone transport. The CSC-CE may attach directly to CEs if it is also configured to be a PE for business VPN services.
- CSC-PE — A PE router managed and operated by the Super Carrier that supports one or more CSC IP VPN services possibly in addition to other traditional PE services.

Figure 9: CSC network topology



25464

In the CSC solution, the CSC-CE and CSC-PE are directly connected by a link that supports MPLS. The CSC-CE distributes an MPLS label for every /32 IPv4 prefix it and any downstream PE uses as a BGP next-hop in routes associated with services offered by the Customer Carrier. BGP must be used as the label distribution protocol between CSC-CE and CSC-PE if the latter device is an SR OS node. Typically, the Customer Carrier and Super Carrier operate as two different Autonomous Systems (ASs) and therefore BGP, more specifically EBGP, is the best label distribution protocol, even if other options are available. The BGP session between CSC-CE and CSC-PE must be single-hop EBGP (or IBGP) if either device is an SR OS node.

In an SR OS CSC-PE, the interface to a CSC-CE is a special type of IP/MPLS interface that belongs to a VPRN configured for CSC mode. This special type of interface is called a CSC VPRN interface throughout the remainder of this chapter. The CSC VPRN interface has many of the same characteristics as a network interface of the base router but its association with a Virtual Routing and Forwarding (VRF) ensures that the traffic and control plane routes of the Customer Carrier are kept separate from other services.

When an SR OS CSC-PE receives a labeled-IPv4 route (with label L1, next-hop N1) from a CSC-CE BGP peer, the following actions take place in the CSC-PE:

1. The BGP route is installed into the routing table of the CSC VPRN (assuming the BGP route is the best route to the destination).
2. If the BGP route matches the VRF export policy, it is advertised to the core Multi-Protocol Border Gateway Protocol (MP-BGP) peers as a VPN-IPv4 route. The advertised label value is changed to label value L2.
3. BGP programs the line cards with an MPLS forwarding entry that swaps label value L2 for L1 and sends the MPLS packet over the CSC VPRN interface associated with next-hop N1.

When an SR OS CSC-PE receives a VPN-IPv4 route (with label L2, next-hop N2) the following actions take place in the CSC-PE:

1. If the VPN-IPv4 route matches the VRF import policy of a CSC VPRN, it is installed into the routing table of that CSC VPRN.
2. If the imported BGP-VPN route matches the BGP export policy associated with a CSC-CE BGP peer, it is advertised to that peer as a labeled-IPv4 route. The advertised label value is changed to label value L3.
3. BGP programs the line cards with an MPLS forwarding entry that swaps label value L3 for L2 and sends the packet inside the MPLS tunnel to next-hop N2.

Once a CSC-CE has learned a labeled-IPv4 route for a remote CSC-CE and vice versa, the two CSC-CEs can set up a BGP session between themselves and exchange VPN routes over this session if they are both PEs with services. Typically, this BGP session will be an IBGP session because the local and remote CSC-CEs belong to the same AS. The Layer 2 VPN and Layer 3 VPN routes exchanged by the CSC-CEs are resolved by the labeled-IPv4 routes they have for each other's /32 IPv4 address.

## Configuration

This section will walk through the steps to configure the CSC solution shown in [Figure 9: CSC network topology](#). The IPv4 addresses in [Figure 9: CSC network topology](#) are the system IP addresses of the routers. The steps are the following:

- Configure CSC-CE-1
- Configure CSC service on CSC-PE-2
- Verify exchange of labeled IPv4 routes between CSC-CE-1 and CSC-PE-2
- Configure core connectivity for CSC-PE-2
- Configure core connectivity for CSC-PE-3
- Configure CSC service on CSC-PE-3
- Verify exchange of VPN-IPv4 routes between CSC-PE-2 and CSC-PE-3
- Configure CSC-CE-4
- Verify exchange of labeled IPv4 routes between CSC-PE-3 and CSC-CE-4
- Configure BGP session between CSC-CE-1 and CSC-CE-4
- Verify exchange of VPN-IPv4 routes between CSC-CE-1 and CSC-CE-4

### Step 1. Configure CSC-CE-1

This example assumes that CSC-CE-1 is a PE router with Layer 2 and Layer 3 VPN services that must extend across the CSC VPN service; assume that there are no further downstream PEs in AS 64496. The configuration of one such Layer 3 VPN service in CSC-CE-1 is as follows:

```
# on CSC-CE-1:
configure
  service
    vprn 1 name "VPRN1" customer 1 create
      interface "loopback-1" create
        address 10.11.30.2/24
        loopback
      exit
    bgp-ipvpn
      mpls
        auto-bind-tunnel
        resolution any
      exit
      route-distinguisher 64496:11
      vrf-target target:64496:1
      no shutdown
    exit
  exit
no shutdown
exit
```

For brevity, the preceding configuration sample omits commands related to SAP IP interfaces, spoke-SDP IP interfaces, PE-CE routing protocols, QoS, IP filters, and so on. The loopback interface is used to test whether this prefix is learned at the remote CSC-CE-4.

The base routing instance of the CSC-CE is configured with the appropriate router ID and autonomous system number and the system interface is configured with an IPv4 address (usually the same as the router ID). If the router ID is not configured, by default, the system IP address is used as the router ID. The interface to CSC-PE-2 is created and configured. The base router configuration of CSC-CE-1 is as follows:

```
# on CSC-CE-1:
configure
  router Base
    interface "int-CSC-CE-1-CSC-PE-2"
      address 192.168.12.1/30
      port 1/1/1:1
      no shutdown
    exit
    interface "system"
      address 192.0.2.1/32
      no shutdown
    exit
    autonomous-system 64496
  exit
```

On CSC-CE-1, BGP is configured as the control plane protocol running on the interface to CSC-PE-2, as follows:

```
# on CSC-CE-1:
configure
  router Base
    bgp
      group "CSC-PE"
        peer-as 64500
        neighbor 192.168.12.2
          family label-ipv4
          export "static-to-BGP"
          split-horizon
        exit
      exit
    no shutdown
  exit
```

The peer type is EBGP (**peer-as** is different from the locally configured **autonomous-system**)

The address family for the EBGP session is **label-ipv4** (the **neighbor** address is an IPv4 address). Family label-IPv4 causes MP-BGP negotiation of the address family for AFI=1 and SAFI=4 (IPv4 NLRI with MPLS labels), as can be observed from the following debug message of the BGP OPEN message (in this example, debugging is enabled on CSC-CE-1 for BGP OPEN messages using the command **debug router bgp open**). This BGP OPEN message can obviously only be seen when the BGP peer is up. The configuration for CSC-PE-2 will be shown later, but in order to have the trace message, it must be configured already.

```
# on CSC-CE-1:
2 2019/05/09 07:38:09.783 UTC MINOR: DEBUG #2001 Base BGP
"BGP: OPEN
Peer 1: 192.168.12.2 - Received BGP OPEN: Version 4
AS Num 64500: Holdtime 90: BGP_ID 192.0.2.2: Opt Length 16 (ExtOpt F)
Opt Para: Type CAPABILITY: Length = 14: Data:
  Cap_Code MP-BGP: Length 4
  Bytes: 0x0 0x1 0x0 0x4
```

```

Cap_Code ROUTE-REFRESH: Length 0
Cap_Code 4-OCTET-ASN: Length 4
Bytes: 0x0 0x0 0xfb 0xf4
"

```

The **split-horizon** command is optional. It prevents a best BGP route from the CSC-PE peer from being re-advertised back to that peer.

The **export** command applies a BGP export policy to the session. The configuration of the export policy on CSC-CE-1 is as follows:

```

# on CSC-CE-1:
configure
  router Base
    policy-options
      begin
        prefix-list "system-IP"
          prefix 192.0.2.1/32 exact
        exit
        policy-statement "static-to-BGP"
          entry 10
            from
              protocol direct
              prefix-list "system-IP"
            exit
            action accept
          exit
        exit
      default-action drop
    exit
  exit
commit

```

The purpose of the BGP export policy is to advertise the system IP address of CSC-CE-1 as a labeled-IPv4 BGP route toward CSC-PE-2.

### Step 2. Configure CSC service on CSC-PE-2

CSC-PE-2 must be configured with a VPRN in **carrier-carrier-vpn** mode to provide CSC service to CSC-CE-1. VPRN 1 is configured on CSC-PE-2, as follows:

```

# on CSC-PE-2:
configure
  service
    vprn 1 name "VPRN1" customer 1 create
      carrier-carrier-vpn
      router-id 192.0.2.2
      autonomous-system 64500
      network-interface "int-CSC-PE-2-CSC-CE-1" create
        address 192.168.12.2/30
        port 1/1/2:1
        no shutdown
      exit
    bgp-ipvpn
      mpls
        auto-bind-tunnel
          resolution any
        exit
        route-distinguisher 64500:12
        vrf-target target:64500:1
        no shutdown
      exit

```

```

exit
  bgp
    group "CSC-CE"
      as-override
      export "BGP-VPN-routes"
      peer-as 64496
      neighbor 192.168.12.1
        family label-ipv4
        split-horizon
      exit
    exit
  no shutdown
exit
no shutdown
exit

```

The **carrier-carrier-vpn** command is mandatory. It cannot be configured if the VPRN currently has any SAP or spoke-SDP access interfaces configured; they must first be disabled if necessary and then deleted.

```

*A:CSC-PE-2>config>service>vprn# carrier-carrier-vpn
INFO: PIP #1195 Cannot toggle carrier-carrier-vpn - service interfaces present

```

The **auto-bind-tunnel** command must be set appropriately for the type of transport desired to other CSC-PEs, but note that GRE is not supported.

```

*A:CSC-PE-2>config>service>vprn>auto-bind-tunnel# resolution-filter gre
MINOR: SVCNMR #1538 auto-bind config not supported - Autobind gre not supported for carrier-
carrier vprn

```

The interface to CSC-CE-1 must be a network interface. A network interface can be associated with an entire Ethernet port, a VLAN sub-interface of an Ethernet port, an entire LAG or a VLAN sub-interface of a LAG. In all cases, the associated Ethernet ports must be configured in network or hybrid mode.

The peer type is EBGp (**peer-as** is different from the locally configured **autonomous-system**).

The address family for the EBGp session is **label-ipv4** (the **neighbor** address is an IPv4 address). Address family label-ipv4 causes MP-BGP negotiation of the address family for AFI=1 and SAFI=4 (IPv4 NLRI with MPLS labels).

The **split-horizon** command is optional. It prevents a best BGP route from the CSC-CE peer from being re-advertised back to that peer.

The **as-override** command replaces CSC-CE-1's AS number (64496) with CSC-PE-2's AS number (64500) in the AS\_PATH attribute of routes advertised to CSC-CE-1. Without this configuration, CSC-CE-1 may reject routes originated by CSC-CE-4 as invalid due to an AS-path loop.

The **export** command applies a BGP export policy to the session. The configuration of the policy is as follows:

```

# on CSC-PE-2:
configure
  router Base
    policy-options
      begin
        policy-statement "BGP-VPN-routes"
          entry 10
            from
              protocol bgp-vpn
            exit
          action accept

```



```

        exit
        exit
        default-action drop
        exit
    exit
    commit
exit
    
```

The effect of the BGP export policy is to re-advertise VPN-IPv4 routes imported into the CSC VPRN (and used for forwarding) to CSC-CE-4.

### Step 3. Verify exchange of labeled IPv4 routes

When steps 1 and 2 have been completed, CSC-CE-1 advertises the labeled-IPv4 route for its system IP address 192.0.2.1/32 to CSC-PE-2. This can be checked in the RIB Out of CSC-CE-1, as follows:

```

*A:CSC-CE-1# show router bgp routes 192.0.2.1/32 label-ipv4 hunt
=====
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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP LABEL-IPV4 Routes
=====
-----
RIB In Entries
-----
-----
RIB Out Entries
-----
Network       : 192.0.2.1/32
Nexthop       : 192.168.12.1
Path Id       : None
To            : 192.168.12.2
Res. Nexthop  : n/a
Local Pref.   : n/a
Aggregator AS : None
Atomic Aggr.  : Not Atomic
AIGP Metric   : None
Connector     : None
Community     : No Community Members
Cluster       : No Cluster Members
Originator Id : None
IPv4 Label    : 524286
Lbl Allocation : NEXT-HOP
Origin        : IGP
AS-Path       : 64496
Route Tag     : 0
Neighbor-AS   : 64496
Orig Validation: NotFound
Source Class  : 0
Interface Name : NotAvailable
Aggregator    : None
MED           : None
IGP Cost      : n/a
Peer Router Id : 192.0.2.2
Label Type    : POP
Dest Class    : 0
-----
Routes : 1
=====
    
```

CSC-CE-1 has advertised a label value of 524286 with the prefix.

The following output shows the received route in the RIB In for VPRN 1 on CSC-PE-2:

```
*A:CSC-PE-2# show router 1 bgp routes 192.0.2.1/32 label-ipv4 hunt
=====
BGP Router ID:192.0.2.2      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 LABEL-IPV4 Routes
=====
-----
RIB In Entries
-----
Network       : 192.0.2.1/32
Nexthop       : 192.168.12.1
Path Id       : None
From          : 192.168.12.1
Res. Nexthop  : 192.168.12.1
Local Pref.   : None
Aggregator AS : None
Atomic Aggr.  : Not Atomic
AIGP Metric   : None
Connector     : None
Community     : No Community Members
Cluster       : No Cluster Members
Originator Id : None
Fwd Class     : None
IPv4 Label   : 524286
Flags         : Used Valid Best IGP In-TTM In-RTM
Route Source  : External
AS-Path       : 64496
Route Tag     : 0
Neighbor-AS   : 64496
Orig Validation: NotFound
Source Class  : 0
Add Paths Send : Default
RIB Priority   : Normal
Last Modified : 00h01m22s
-----
Interface Name : int-CSC-PE-2-CSC-CE-1
Aggregator     : None
MED            : None
IGP Cost       : 0
Peer Router Id : 192.0.2.1
Priority        : None
-----
RIB Out Entries
-----
-----
Routes : 1
=====
```

#### Step 4. Configure core connectivity for CSC-PE-2

The next step is to configure the base router instance of CSC-PE-2 so that it can exchange VPN-IPv4 routes with CSC-PE-3 (and potentially other CSC-PEs). This requires:

- Router ID and autonomous system configuration.
- Network interface creation and configuration, including assignment of an IPv4 address to the system interface.
- Configuration of the IGP protocol; in this example, IS-IS is used.
- Configuration of the LDP protocol (optional).

- Configuration of RSVP LSPs used to reach remote CSC-PE devices (optional).
- Configuration of the BGP protocol.

The base router configuration of CSC-PE-2 is as follows:

```
# on CSC-PE-2
configure
router Base
  interface "int-CSC-PE-2-CSC-PE-3"
    address 192.168.23.1/30
    port 1/1/1:1000
    no shutdown
  exit
  interface "system"
    address 192.0.2.2/32
    no shutdown
  exit
  autonomous-system 64500
  isis 0
    level-capability level-2
    area-id 49.01
    level 2
      wide-metrics-only
    exit
  interface "system"
    passive
    no shutdown
  exit
  interface "int-CSC-PE-2-CSC-PE-3"
    interface-type point-to-point
    no shutdown
  exit
  no shutdown
exit
ldp
  interface-parameters
    interface "int-CSC-PE-2-CSC-PE-3" dual-stack
      ipv4
        no shutdown
      exit
      no shutdown
    exit
  exit
  targeted-session
  exit
  no shutdown
exit
bgp
  group "core"
    type internal
    neighbor 192.0.2.3
      family vpn-ipv4
    exit
  exit
  no shutdown
exit
```

The peer type is IBGP (**type internal**). It is also possible to configure this in a similar way as for eBGP, with the same value for **peer-as** as the locally configured **autonomous-system**).

The transport for the IBGP session is IPv4 (the **neighbor** address is an IPv4 address).

The **family vpn-ipv4** command causes MP-BGP negotiation of the address family for AFI=1 and SAFI=128 (=0x80), as can be observed from the following debug trace of the OPEN message from CSC-PE-2 to CSC-PE-3.

```

1 2022/04/05 07:59:07.576 UTC MINOR: DEBUG #2001 Base BGP
"BGP: OPEN
Peer 1: 192.0.2.3 - Send (Passive) BGP OPEN: Version 4
AS Num 64500: Holdtime 90: BGP_ID 192.0.2.2: Opt Length 20 (ExtOpt F)
Opt Para: Type CAPABILITY: Length = 18: Data:
  Cap_Code GRACEFUL-RESTART: Length 2
  Bytes: 0x0 0x78
  Cap_Code MP-BGP: Length 4
  Bytes: 0x0 0x1 0x0 0x80
  Cap_Code ROUTE-REFRESH: Length 0
  Cap_Code 4-OCTET-ASN: Length 4
  Bytes: 0x0 0x0 0xfb 0xf4
"

```

### Step 5. Configure core connectivity for CSC-PE-3

The next step is to configure the base router instance of CSC-PE-3 so that it can exchange VPN-IPv4 routes with CSC-PE-2 and potentially other CSC-PEs. This requires:

- Router ID and AS configuration.
- Network interface creation and configuration, including assignment of an IPv4 address to the system interface.
- Configuration of the IGP protocol; in this example IS-IS is used.
- Configuration of the LDP protocol (optional).
- Configuration of RSVP LSPs used to reach remote CSC-PE devices (optional).
- Configuration of the BGP protocol.

The base router configuration of CSC-PE-3 is as follows:

```

# on CSC-PE-3
configure
router Base
  interface "int-CSC-PE-3-CSC-PE-2"
    address 192.168.23.2/30
    port 1/1/2:1000
    no shutdown
  exit
  interface "system"
    address 192.0.2.3/32
    no shutdown
  exit
  autonomous-system 64500
  isis 0
    level-capability level-2
    area-id 49.01
    level 2
      wide-metrics-only
    exit
  interface "system"
    passive
    no shutdown
  exit
  interface "int-CSC-PE-3-CSC-PE-2"
    interface-type point-to-point
    no shutdown

```

```

        exit
        no shutdown
    exit
    ldp
        interface-parameters
            interface "int-CSC-PE-3-CSC-PE-2" dual-stack
                ipv4
                    no shutdown
                exit
                no shutdown
            exit
        exit
        targeted-session
        exit
        no shutdown
    exit
    bgp
        group "core"
            type internal
            cluster 192.0.2.3
            neighbor 192.0.2.2
                family vpn-ipv4
                    split-horizon
            exit
        exit
        no shutdown
    exit

```

The peer type is IBGP (**type internal**. Can also be configured with **peer-as** equal to the locally configured **autonomous-system**).

The transport for the IBGP session is IPv4 (the **neighbor** address is an IPv4 address).

The **family vpn-ipv4** command causes MP-BGP negotiation of the address family for AFI=1 and SAFI=128.

The **cluster** command configures CSC-PE-2 as a route reflector for clients in the BGP group "core". This is not required and in a more typical deployment, the route reflector would be a separate router from any CSC-PE.

#### Step 6. Configure CSC service on CSC-PE-3

CSC-PE-3 must be configured with a VPRN in **carrier-carrier-vpn** mode to provide CSC service to CSC-CE-4. The configuration of the VPRN is as follows:

```

# on CSC-PE-3:
configure
    service
        vprn 1 name "VPRN1" customer 1 create
            carrier-carrier-vpn
            router-id 192.0.2.3
            autonomous-system 64500
            network-interface "int-CSC-PE-3-CSC-CE-4" create
                address 192.168.34.1/30
                port 1/1/1:1
                no shutdown
            exit
        bgp-ipvpn
            mpls
                auto-bind-tunnel
                resolution any
            exit
            route-distinguisher 64500:13

```

```

        vrf-target target:64500:1
        no shutdown
    exit
exit
bgp
    group "CSC-CE"
        as-override
        export "BGP-VPN-routes"
        peer-as 64496
        neighbor 192.168.34.2
            family label-ipv4
            split-horizon
        exit
    exit
    no shutdown
exit
    no shutdown
exit

```

The **carrier-carrier-vpn** command is mandatory. It cannot be configured if the VPRN has any SAP or spoke-SDP access interfaces configured; they must first be disabled if necessary and removed.

The **auto-bind-tunnel** command must be set appropriately for the type of transport desired to other CSC-PEs, but GRE is not supported.

The interface to CSC-CE-4 must be a network interface. A network interface can be associated with an entire Ethernet port, a VLAN sub-interface of an Ethernet port, an entire LAG or a VLAN sub-interface of a LAG. In all cases, the associated Ethernet ports must be configured in network or hybrid mode.

The peer type is EBGp (**peer-as** is different from the locally configured **autonomous-system**).

The address family for the EBGp session is **label-ipv4** (the **neighbor** address is an IPv4 address). Address family label-ipv4 causes MP-BGP negotiation of the address family for AFI=1 and SAFI=4 (IPv4 NLRI with MPLS labels).

The **split-horizon** command is optional. It prevents a best BGP route from the CSC-CE peer from being re-advertised back to that peer.

The **as-override** command replaces CSC-CE-4's AS number 64496 with CSC-PE-3's AS number 64500 in the AS\_PATH attribute of routes advertised to CSC-CE-4. Without this configuration, CSC-CE-4 may reject routes originated by CSC-CE-1 as invalid due to an AS-path loop.

The **export** command applies a BGP export policy to the session. The configuration of the policy is as follows:

```

# on CSC-PE-3:
configure
    router Base
        policy-options
            begin
            policy-statement "BGP-VPN-routes"
                entry 10
                    from
                        protocol bgp-vpn
                    exit
                    action accept
                exit
            exit
            default-action drop
        exit
    exit
commit

```

exit

The effect of the BGP export policy is to re-advertise VPN-IPv4 routes imported into the CSC VPRN (and used for forwarding) to CSC-CE-4.

**Step 7.** Verify exchange of VPN-IPv4 routes between CSC-PE-2 and CSC-PE-3.

When the preceding steps have been completed, CSC-PE-2 advertises the labeled-IPv4 route for 192.0.2.1/32 (the system IP address of CSC-CE-1) to CSC-PE-3. This can be checked in the RIB Out of CSC-PE-2, as follows:

```
*A:CSC-PE-2# show router bgp routes 192.0.2.1/32 vpn-ipv4 hunt
=====
BGP Router ID:192.0.2.2      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
=====
RIB In Entries
-----
RIB Out Entries
-----
Network       : 192.0.2.1/32
Nexthop       : 192.0.2.2
Route Dist.   : 64500:12      VPN Label     : 524284
Path Id       : None
To            : 192.0.2.3
Res. Nexthop  : n/a
Local Pref.   : 100
Aggregator AS : None          Interface Name : NotAvailable
Atomic Aggr.  : Not Atomic   Aggregator    : None
AIGP Metric   : None          MED           : None
Connector     : None          IGP Cost      : n/a
Community     : target:64500:1
Cluster       : No Cluster Members
Originator Id : None          Peer Router Id : 192.0.2.3
Origin        : IGP
AS-Path       : 64496
Route Tag     : 0
Neighbor-AS   : 64496
Orig Validation: N/A
Source Class  : 0             Dest Class    : 0
-----
Routes : 1
=====
```

CSC-PE-2 has advertised a VPN label value of 524284 with the prefix.

The following output shows the received route in the RIB In of CSC-PE-3:

```
*A:CSC-PE-3# show router bgp routes 192.0.2.1/32 vpn-ipv4 hunt
=====
BGP Router ID:192.0.2.3      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
=====
-----
RIB In Entries
-----
Network       : 192.0.2.1/32
NextHop       : 192.0.2.2
Route Dist.   : 64500:12          VPN Label      : 524284
Path Id       : None
From          : 192.0.2.2
Res. NextHop  : n/a
Local Pref.   : 100
Aggregator AS : None             Interface Name : int-CSC-PE-3-CSC-PE-2
Atomic Aggr.  : Not Atomic       Aggregator     : None
AIGP Metric   : None             MED            : None
Connector     : None             IGP Cost       : 10
Community     : target:64500:1
Cluster       : No Cluster Members
Originator Id : None             Peer Router Id  : 192.0.2.2
Fwd Class     : None             Priority        : None
Flags         : Used Valid Best IGP
Route Source  : Internal
AS-Path       : 64496
Route Tag     : 0
Neighbor-AS   : 64496
Orig Validation: N/A
Source Class  : 0                 Dest Class     : 0
Add Paths Send : Default
Last Modified  : 00h00m25s
VPRN Imported  : 1

-----
RIB Out Entries
-----
-----
Routes : 1
=====

```

The label swap entries that BGP programmed in the line cards of CSC-PE-2 based on the received labeled-IPv4 route from CSC-CE-1 (Label Origin = ExtCarCarVpn) and the advertised VPN-IPv4 route to CSC-PE-3, as follows:

```

*A:CSC-PE-2# show router bgp inter-as-label

=====
BGP Inter-AS labels
Flags: B - entry has backup, P - entry is promoted
=====
NextHop           Received      Advertised     Label
                  Label         Label          Origin
-----
192.168.12.1     524286       524284        ExtCarCarVpn
-----
Total Labels allocated: 1
=====

```



### Step 8. Configure CSC-CE-4

In this example, CSC-CE-4 is a PE router with Layer 2 and Layer 3 VPN services that must extend across the CSC VPN service. The configuration of one such Layer 3 VPN service in CSC-CE-4 is as follows:

```
# on CSC-CE-4
configure
  service
    vprn 1 name "VPRN1" customer 1 create
    interface loopback-1 create
      address 10.14.30.2/24
      loopback
    exit
    bgp-ipvpn
      mpls
        auto-bind-tunnel
        resolution any
      exit
      route-distinguisher 64496:14
      vrf-target target:64496:1
      no shutdown
    exit
  exit
  no shutdown
exit
```

For brevity, the preceding configuration sample omits commands related to SAP IP interfaces, spoke-SDP IP interfaces, PE-CE routing protocols, QoS, IP filters, and so on.

The base routing instance of CSC-CE-4 is configured with the appropriate router ID and AS number and the system interface has an IPv4 address (usually the same as the router ID). The interface to CSC-PE-3 is configured. The base router configuration of CSC-CE-4 is as follows:

```
# on CSC-CE-4
configure
  router Base
    interface "int-CSC-CE-4-CSC-PE-3"
      address 192.168.34.2/30
      port 1/1/2:1
      no shutdown
    exit
    interface "system"
      address 192.0.2.4/32
      no shutdown
    exit
    autonomous-system 64496
  exit
```

BGP is configured as the control plane protocol running on the interface to CSC-PE-3, as follows:

```
# on CSC-CE-4
configure
  router Base
    bgp
      group "CSC-PE"
        peer-as 64500
        neighbor 192.168.34.1
          family label-ipv4
          export "static-to-BGP"
          split-horizon
        exit
      exit
    exit
```

```
no shutdown
exit
```

The peer type is EBGp (**peer-as** is different from the locally configured **autonomous-system**).

The address family for the EBGp session is **label-ipv4** (the **neighbor** address is an IPv4 address). Address family label-ipv4 causes MP-BGP negotiation of the address family for AFI=1 and SAFI=4 (IPv4 NLRI with MPLS labels).

The **split-horizon** command is optional. It prevents a best BGP route from the CSC-PE peer from being re-advertised back to that peer.

The **export** command applies a BGP export policy to the session. The configuration of the policy is as follows:

```
# on CSC-CE-4
configure
router Base
  policy-options
  begin
  prefix-list "system-IP"
    prefix 192.0.2.4/32 exact
  exit
  policy-statement "static-to-BGP"
    entry 10
      from
        protocol direct
        prefix-list "system-IP"
      exit
      action accept
    exit
  exit
  default-action drop
  exit
exit
commit
exit
```

The purpose of the BGP export policy is to advertise the system IP address of CSC-CE-4 as a labeled-IPv4 BGP route toward CSC-PE-3.

**Step 9.** Verify exchange of labeled IPv4 routes between CSC-PE-3 and CSC-CE-4

When the preceding steps are completed, CSC-PE-3 advertises the labeled-IPv4 route for 192.0.2.1/32 to CSC-CE-4. This can be checked in the RIB Out for VPRN 1 on CSC-PE-3, as follows:

```
*A:CSC-PE-3# show router 1 bgp routes 192.0.2.1/32 label-ipv4 hunt
=====
BGP Router ID:192.0.2.3      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 LABEL-IPV4 Routes
=====
-----
RIB In Entries
-----
```

```

-----
RIB Out Entries
-----
Network       : 192.0.2.1/32
Nextthop     : 192.168.34.1
Path Id      : None
To           : 192.168.34.2
Res. Nextthop : n/a
Local Pref.  : n/a
Aggregator AS : None
Atomic Aggr. : Not Atomic
AIGP Metric  : None
Connector    : None
Community    : No Community Members
Cluster      : No Cluster Members
Originator Id : None
IPv4 Label   : 524284
Lbl Allocation : NEXT-HOP
Origin       : IGP
AS-Path      : 64500 64500
Route Tag    : 0
Neighbor-AS  : 64500
Orig Validation: NotFound
Source Class : 0
Interface Name : NotAvailable
Aggregator    : None
MED           : None
IGP Cost      : n/a
Peer Router Id : 192.0.2.4
Label Type    : SWAP
Dest Class    : 0

-----
Routes : 1
=====

```

CSC-PE-3 has advertised a label value of 524284 with the prefix.

The following output shows the received route in the RIB In of CSC-CE-4:

```

*A:CSC-CE-4# show router bgp routes 192.0.2.1/32 label-ipv4 hunt
=====
BGP Router ID:192.0.2.4      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
Origin codes  : i - IGP, e - EGP, ? - incomplete

=====
BGP Routes
=====
-----
RIB In Entries
-----
Network       : 192.0.2.1/32
Nextthop     : 192.168.34.1
Path Id      : None
From         : 192.168.34.1
Res. Nextthop : 192.168.34.1
Local Pref.  : None
Aggregator AS : None
Atomic Aggr. : Not Atomic
AIGP Metric  : None
Connector    : None
Community    : target:64500:1
Cluster      : No Cluster Members
Originator Id : None
Fwd Class    : None
IPv4 Label   : 524284
Flags        : Used Valid Best IGP
Interface Name : int-CSC-CE-4-CSC-PE-3
Aggregator    : None
MED           : None
Peer Router Id : 192.0.2.3
Priority       : None

```

```
Route Source : External
AS-Path      : 64500 64500
Route Tag    : 0
Neighbor-AS  : 64500
Orig Validation: NotFound
Source Class : 0
Dest Class   : 0
Add Paths Send : Default
Last Modified : 00h00m53s
```

```
-----
RIB Out Entries
-----
```

```
Routes : 1
=====
```

The BGP distributed labels are programmed in the line cards of CSC-PE-3 based on the received VPN-IPv4 routes from CSC-PE-2 (Label Origin = Internal) and the advertised labeled-IPv4 routes to CSC-CE-4:

```
*A:CSC-PE-3# show router 1 bgp inter-as-label
```

```
=====
BGP Inter-AS labels
Flags: B - entry has backup, P - entry is promoted
=====
NextHop                Received      Advertised    Label
                        Label         Label         Origin
-----
192.0.2.2              524284       524284       Internal
192.0.2.2              524285       524286       Internal
-----
Total Labels allocated: 2
=====
```

In the preceding output, the first entry for NextHop 192.0.2.2 corresponds to the prefix 192.0.2.1/32; recall from Step 7 that CSC-PE-3 received the VPN-IPv4 route with label value 524284 and it can be seen from this step that it re-advertised the route to CSC-CE-4 with the same label value 524284.

#### Step 10. Configure BGP session between CSC-CE-1 and CSC-CE-4

The final step in the setup of the CSC solution shown in [Figure 9: CSC network topology](#) is the creation of a BGP session between CSC-CE-1 and CSC-CE-4 so that they can exchange routes belonging to VPN services they support. The configuration of this BGP session on CSC-CE-1 is as follows:

```
# on CSC-CE-1:
configure
  router Base
  bgp
    group "CSC-CE"
      type internal
      neighbor 192.0.2.4
        family vpn-ipv4
    exit
  exit
  no shutdown
exit
```

The configuration of the BGP session on CSC-CE-4 is similar, as follows:

```
# on CSC-CE-4:
configure
```

```

router Base
  bgp
    group "CSC-CE"
      type internal
      neighbor 192.0.2.1
        family vpn-ipv4
      exit
    exit
  no shutdown
exit

```

The configuration of the BGP session between CSC-CE-1 and CSC-CE-4 has the following properties:

- The peer type is IBGP (**type internal**. Alternatively, **peer-as** can be configured with the same value as the locally configured **autonomous-system**).
- The transport for the IBGP session is IPv4 (the **neighbor** address is an IPv4 address).
- The **family vpn-ipv4** command causes MP-BGP negotiation of the address family for AFI=1 and SAFI=128.

#### Step 11. Verify exchange of VPN-IPv4 routes

When the preceding steps have been completed, CSC-PE-3 can advertise a VPN-IPv4 route for some IP prefix (for example, 10.11.30.0/24) to CSC-CE-4. This can be checked in the RIB In of CSC-CE-4 as follows:

```

*A:CSC-CE-4# show router bgp routes 10.11.30.0/24 vpn-ipv4 hunt
=====
BGP Router ID:192.0.2.4      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP VPN-IPv4 Routes
=====
RIB In Entries
-----
Network       : 10.11.30.0/24
NextHop       : 192.0.2.1
Route Dist.   : 64496:11      VPN Label     : 524287
Path Id       : None
From          : 192.0.2.1
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     : target:64496:1
Cluster       : No Cluster Members
Originator Id : None         Peer Router Id : 192.0.2.1
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  : 00h00m45s
VPRN Imported  : 1
```

```
-----
RIB Out Entries
-----
```

```
Routes : 1
=====
```

The following command can be used to check that CSC-CE-4 has properly installed the preceding VPN-IPv4 route into the routing table of the importing VPRN service:

```
*A:CSC-CE-4# show router 1 route-table
```

```
=====
Route Table (Service: 1)
=====
```

Dest Prefix[Flags] Next Hop[Interface Name]	Type	Proto	Age Metric	Pref
<b>10.11.30.0/24</b>	<b>Remote</b>	<b>BGP VPN</b>	<b>00h01m56s</b>	<b>170</b>
<b>192.0.2.1 (tunneled:BGP)</b>			<b>1000</b>	
10.14.30.0/24	Local	Local	00h04m34s	0
loopback-1			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

Carrier Supporting Carrier is a scalable and secure solution for using an infrastructure IP VPN to transport traffic between dispersed CSC-CE devices belonging to an ISP or other service provider. Many different topology models are supported by SR OS. This chapter has explored one simplified configuration that can serve as the basis for more complicated setups.

# Flexible Algorithms for SRv6-based VPRNs

This chapter provides information about flexible algorithms (Flex-Algorithm) for VPRNs that are based on segment routing over IPv6 (SRv6).

Topics in this chapter include:

- [Applicability](#)
- [Overview](#)
- [Configuration](#)
- [Conclusion](#)

## Applicability

The information and configuration in this chapter are based on SR OS Release 22.10.R1. The Flex-Algorithm for SRv6-based VPRNs feature is supported on FP-based platforms with FP4-based network ports in SR OS Release 21.5.R2 and later.

## Overview

The Flex-Algorithm for SRv6-based VPRNs feature allows the computation of constraint-based paths across an SRv6-enabled network, based on metrics other than the default interior gateway protocol (IGP) metrics. The supported metrics are:

- IGP metrics
- link delay metrics
- traffic engineering (TE) metrics

Based on the metrics that are specifically configured for it, each Flex-Algorithm instance computes optimum paths across routers that participate in the Flex-Algorithm instance. For these paths, the IGP protocol automatically creates SRv6 tunnels between every pair of routers participating in the Flex-Algorithm instance. Two or more routers participate in a single Flex-Algorithm instance; a single router may participate in multiple Flex-Algorithm instances.

At least one router advertises (via extensions to the IGP protocol) the flexible algorithm definition (FAD). The **metric-type** *{igp|te-metric|delay}* command in the **router Base flexible-algorithm-definitions flex-algo <fad-name>** context configures the metric type that the Flex-Algorithm instance uses: *igp*, *delay*, or *te-metric*. The router that advertises the FAD typically also participates in the Flex-Algorithm instance. The other routers participate in the advertised Flex-Algorithm instance, without also advertising it. For reasons of redundancy, multiple routers may advertise the same FAD. In that case, the configuration of that FAD should be identical on all these routers. If not, all routers that participate in the Flex-Algorithm instance install from conflicting FADs only the FAD that has the highest priority value. If conflicting FADs have the same priority value, all routers that participate in the Flex-Algorithm instance install only the FAD that is advertised by the IS-IS-enabled router with the highest IS-IS system ID (or by the OSPF-enabled router with the highest OSPF router ID).

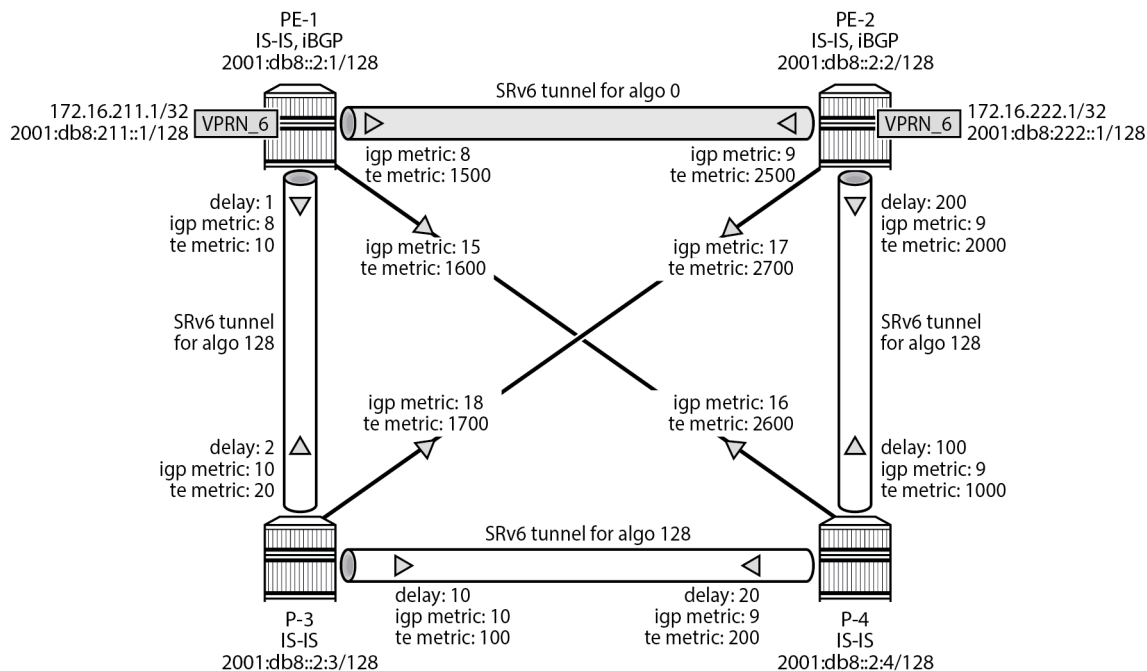
The **algorithm <flex-algo-id>** command in the **router Base segment-routing segment-routing-v6 locator <locator-name>** context, associates each Flex-Algorithm instance (algorithm 128 to algorithm 255) with one specific SRv6 locator, which must be different from the base algorithm (algorithm 0) SRv6 locator. This algorithm identifier is included in the SRv6 locator TLV when advertising the locator into IS-IS. The same FAD may be used in multiple Flex-Algorithm instances. Each Flex-Algorithm instance is associated with, at most, one SRv6 locator. Each SRv6 locator has, at most, one associated Flex-Algorithm instance. The **default-locator <name>** command in the **service vprn <service-id> bgp-ipvpn segment-routing-v6 srv6-instance <[1..2]>** context configures the SRv6 locator that the VPRN data traffic across SRv6-enabled networks uses.

The further processing of the Flex-Algorithm-based VPRN data traffic across SRv6-enabled networks follows that of the base algorithm-based VPRN data traffic across SRv6-enabled networks, as described in the "Segment Routing over IPv6 for VPRN" chapter in the *7750 SR and 7950 XRS Segment Routing and PCE Advanced Configuration Guide for Classic CLI*.

## Configuration

**Figure 10: Example topology** shows the example topology with four routers. The SRv6-enabled network that it represents comprises PE-1 and PE-2 in the control and data plane, and P-3 and P-4 in the data plane only. The SRv6-enabled network has only IPv6 addresses and interfaces. IS-IS is configured on all routers. BGP is configured only on PE-1 and PE-2.

Figure 10: Example topology



38300

PE-1 and PE-2 are SRv6-enabled routers that each contain a VPRN instance. In this example, bidirectional IPv4 and IPv6 VPRN traffic flows are enabled between PE-1 and PE-2.



To illustrate what IS-IS interface metrics are used and how, the IS-IS interface metrics are configured explicitly (that is, differently from their default values), using the **metric** *<ipv4-metric>* command (for IPv4 unicast traffic) and the **ipv6-unicast-metric** *<ipv6-metric>* command (for IPv6 unicast traffic) in the **router Base isis 0 interface <ip-int-name> level {1|2}** context. Different values can be applied for IS-IS level 1 and IS-IS level 2; for each IS-IS level, a distinction can be made between IPv4 unicast traffic and IPv6 unicast traffic. For each IS-IS level and traffic type, different values can be configured in the two directions between two routers.

As a first example, the link delay metric is chosen for the Flex-Algorithm operation. The link delay metrics are configured explicitly, using the **static** *<value>* command (value in microseconds) in the **router Base interface <interface-name> if-attribute delay** context, only on the links between PE-1 and P-3, P-3 and P-4, and P-4 and PE-2. These metrics are configured differently in the two directions on each link (shown in the example topology as: "delay: *<value in microseconds>*"). The link delay metric can be configured on links between any pair of routers participating in the Flex-Algorithm instance. Any link that does not have the link delay metric configured is excluded from the Flex-Algorithm instance computation, which may result in no valid path between the ingress and egress routers. The link delay metric values are used for both IPv4 unicast traffic and IPv6 unicast traffic.

As a second example, the TE metric is chosen for the Flex-Algorithm operation. The TE metrics are configured explicitly, using the **te-metric** *<value>* command in the **router Base mpls interface <interface-name>** context, on all links. These metrics are configured differently in the two directions on each link, such that the direct link between the two routers PE-1 and PE-2 is always preferred.

The **ping** and **traceroute** commands between IPv4 and IPv6 loopback addresses in the VPRNs, as described in following sections, are used to simulate data traffic.

SRv6 requires wide metrics to match the 32-bit metric field in SRv6 locator TLV. The example configuration has wide metrics configured only for level 2. So, only the explicitly configured IS-IS level 2 interface metric values are used. Also, multi-topology and multi-protocol are not enabled in the example configuration. So, the explicitly configured IS-IS level 2 interface metric values are used for both IPv4 unicast traffic and IPv6 unicast traffic.

## Configure the router

This configuration includes:

- ports and IPv6-only interfaces on PE-1, PE-2, P-3, and P-4, with link delay metrics configured where needed
- 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 in the *7750 SR and 7950 XRS Segment Routing and PCE Advanced Configuration Guide for Classic CLI*
- IS-IS on PE-1, PE-2, P-3, and P-4, which includes:
  - level 2 capability with wide metrics, and IPv4 metrics on all level 2 IS-IS interfaces
  - native IPv6 routing
  - the **traffic-engineering** and **traffic-engineering-options** commands, as a best practice to advertise the router capability within the autonomous system (AS)
- BGP on PE-1 and PE-2, with internal group "gr\_v6\_internal", which includes:
  - the IPv4 and IPv6 families
  - **extended-nh-encoding** for IPv4

- **advertise-ipv6-next-hops** for IPv4
- BGP neighbor **system** IPv6 addresses
- **next-hop-self**

As the core network topology uses IPv6 for BGP peering (with IPv6 next hop addresses), the commands **advertise-ipv6-next-hops** and **extended-nh-encoding** need to be applied at the BGP, group, or neighbor level, so as to advertise and receive IPv4 routes with IPv6 next hop addresses. 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. A similar configuration applies for PE-2. P-3 and P-4 have no BGP configuration.

```
*A: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::168:12:1/126
    exit
  no shutdown
  exit
  interface "int-PE-1-P-3"
    description "interface between PE-1 and P-3"
    port 1/1/c2/1:1000
    ipv6
      address 2001:db8::168:13:1/126
    exit
    if-attribute
      delay
        static 1 # microseconds
      exit
    exit
  no shutdown
  exit
  interface "int-PE-1-P-4"
    description "interface between PE-1 and P-4"
    port 1/1/c3/1:1000
    ipv6
      address 2001:db8::168:14:1/126
    exit
  no shutdown
  exit
  interface "system"
    description "system interface of PE-1"
    ipv6
      address 2001:db8::2:1/128
    exit
  no shutdown
  exit
  isis 0
    router-id 1.1.1.1
    level-capability level-2
    area-id 49.0001
    traffic-engineering
    traffic-engineering-options
      ipv6
        application-link-attributes
```

```
        exit
    exit
    advertise-router-capability as
    ipv6-routing native
    level 2
        wide-metrics-only    # required for SRv6
    exit
    interface "system"
        passive
        no shutdown
    exit
    interface "int-PE-1-PE-2"
        interface-type point-to-point
        level 2
            metric 8
        exit
        no shutdown
    exit
    interface "int-PE-1-P-3"
        interface-type point-to-point
        level 2
            metric 8
        exit
        no shutdown
    exit
    interface "int-PE-1-P-4"
        interface-type point-to-point
        level 2
            metric 15
        exit
        no shutdown
    exit
    no shutdown
exit
bgp
    min-route-advertisement 1
    router-id 1.1.1.10
    rapid-withdrawal
    split-horizon
    group "gr_v6_internal"
        description "internal bgp group on PE-1"
        family ipv4 ipv6
        next-hop-self
        type internal
        extended-nh-encoding ipv4
        advertise-ipv6-next-hops ipv4
        neighbor 2001:db8::2:2
        exit
    exit
    no shutdown
exit
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", which includes:
  - 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 the VPRN on PE-1. A similar configuration applies for the VPRN on PE-2.

```
*A:PE-1# configure service
  vprn 6 name "VPRN_6" customer 1 create
    description "VPRN_6 on PE-1"
    autonomous-system 64500
    interface "lb_itf_vprn" create
      address 172.16.211.1/32
      description "VPRN_6 interface on PE-1 for external subnet"
      ipv6
        address 2001:db8:211::1/128
      exit
    loopback
  exit
  bgp
    group "gr_v6_vprn"
      description "external bgp group for VPRN_6 on PE-1"
      family ipv4 ipv6
      extended-nh-encoding ipv4
      advertise-ipv6-next-hops ipv4
      neighbor 2001:db8:101::1
        type external
        peer-as 64501
      exit
    exit
  no shutdown
exit
no shutdown
exit all
```

## Configure SRv6 in the router Base context on PE-1 and PE-2

Configure the SRv6 **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 SRv6 **locator "PE-1\_loc"**.

```
*A:PE-2# configure router Base segment-routing segment-routing-v6
  locator "PE-2_loc"
    block-length 48
    function-length 20
    no algorithm # algo 0
    prefix
      ip-prefix 2001:db8:aaaa:102::/64
    exit
  no shutdown
exit all
```

Configure the FPEs on PE-1 and PE-2.

```
*A:PE-2# configure
  fwd-path-ext
    fpe 1 create
```

```

    path pxc 1
    srv6 origination
        interface-a
        exit
        interface-b
        exit
    exit
exit
fpe 2 create
    path pxc 2
    srv6 termination
        interface-a
        exit
        interface-b
        exit
    exit
exit
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 <locator-name>** context on PE-2. Perform a similar configuration on PE-1, for SRv6 locator "PE-1\_loc". For more information, see the "Segment Routing over IPv6" chapter in the *7750 SR and 7950 XRS Segment Routing and PCE Advanced Configuration Guide for Classic CLI*.

```

*A:PE-2# configure router Base segment-routing
    segment-routing-v6
        origination-fpe 1
        locator "PE-2_loc"
            termination-fpe 2
            no shutdown
        exit
    exit all

```

Configure the SRv6 End function (equivalent to an IPv4 node SID) in the **router Base segment-routing segment-routing-v6 base-routing-instance locator <locator-name>** context on PE-2. Perform a similar configuration on PE-1, for SRv6 locator "PE-1\_loc".

```

*A:PE-2# configure router Base segment-routing segment-routing-v6 base-routing-instance
    locator "PE-2_loc"
        function
            end 1
            srh-mode usp
        exit
    exit
    exit all

```

Advertise the SRv6 locator in IS-IS while ensuring level 2 capability on PE-2. Perform a similar configuration on PE-1, for SRv6 locator "PE-1\_loc".

```

*A:PE-2# configure router Base isis 0
    segment-routing-v6
        locator "PE-2_loc"
            level-capability level-2
            level 2
        exit
    exit
    no shutdown
    exit all

```

## Configure SRv6 for the VPRNs on PE-1 and PE-2

On PE-1 and PE-2, extend the BGP advertisements to include the VPN-IPv4 and VPN-IPv6 families.

```
*A:PE-1#/*A:PE-2# configure router Base
  bgp
    rapid-update vpn-ipv4 vpn-ipv6
    group "gr_v6_internal"
      family ipv4 ipv6 vpn-ipv4 vpn-ipv6
      extended-nh-encoding ipv4 vpn-ipv4
      advertise-ipv6-next-hops ipv4 vpn-ipv4 vpn-ipv6
    exit
  no shutdown
exit all
```

On PE-2, create an SRv6 instance for the VPRN service. Use the SRv6 locator from the **router Base segment-routing segment-routing-v6 <instance>** context and configure End.DT4 and End.DT6 functions for it.

Use the created SRv6 instance in the **service vprn <service-id> bgp-ipvpn segment-routing-v6** context, with the configured SRv6 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 SRv6 locator as the default locator, the PE-1 system IPv6 address as the source address, and a different route distinguisher.

```
*A:PE-2# configure service
  vprn 6
    segment-routing-v6 1 create
      locator "PE-2_loc"
      function
        end-dt4
        end-dt6
      exit
    exit
  exit
  bgp-ipvpn
    segment-routing-v6
      route-distinguisher 192.0.2.2:6
      srv6-instance 1 default-locator "PE-2_loc"
      source-address 2001:db8::2:2
      vrf-target target:64506:6
      no shutdown
    exit
  exit
  no shutdown
exit all
```

## Verify data traffic

At this point, using **ping** and **traceroute** commands, verify that IPv4 data traffic is possible between the local VPRN on PE-1 and the remote VPRN on PE-2:

```
*A:PE-1# ping router 6 172.16.222.1
PING 172.16.222.1 56 data bytes
64 bytes from 172.16.222.1: icmp_seq=1 ttl=64 time=1.49ms.
---snip---
```

```

---- 172.16.222.1 PING Statistics ----
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min = 1.29ms, avg = 1.47ms, max = 1.65ms, stddev = 0.116ms

*A:PE-1# traceroute router 6 172.16.222.1
traceroute to 172.16.222.1, 30 hops max, 40 byte packets
 1 172.16.222.1 (172.16.222.1)  1.64 ms  1.71 ms  1.63 ms
    
```

This data traffic uses the SRv6 tunnels over the direct link between PE-1 and PE-2.

For IPv4 data traffic, the VPRN routing table shows the next hop and the applicable IGP metric for the route to 172.16.222.1.

```

*A:PE-1# show router 6 route-table 172.16.222.1

=====
Route Table (Service: 6)
=====
Dest Prefix[Flags]                Type   Proto   Age      Pref
  Next Hop[Interface Name]                Metric
-----
172.16.222.1/32                    Remote BGP VPN 00h01m52s 170
  2001:db8:aaaa:102:8000:: (tunneled:SRV6)      8
-----
No. of Routes: 1
---snip---
=====
    
```

The next hop is the End.DT4 SRv6 SID of the SRv6 locator "PE-2\_loc" for the VPRN on PE-2, which PE-1 learns from a BGP update from PE-2. The SRv6 tunnel to this next hop has label 524288.

```

*A: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:102:7fff:f000::     End.DT6   524287
  PE-2_loc
  SvcId: 6 Name: VPRN_6
2001:db8:aaaa:102:8000::         End.DT4   524288
  PE-2_loc
  SvcId: 6 Name: VPRN_6
-----
SIDs : 3
=====
    
```

The applicable IGP metric is 8, which corresponds with the IS-IS level 2 "int-PE-1-PE-2" interface metric value.

```

*A:PE-1# show router isis 0 interface

=====
Rtr Base ISIS Instance 0 Interfaces
    
```

```

=====
Interface                               Level CircID Oper   L1/L2 Metric   Type
                               State
-----
system                                 L1L2  1      Up     0/0
int-PE-1-PE-2                       L1L2  2      Up     6/8
int-PE-1-P-3                          L1L2  3      Up     6/8
int-PE-1-P-4                          L1L2  4      Up    11/15
-----
Interfaces : 4
=====

```

The **show router isis 0 topology** command lists the IS-IS nodes, and for each IS-IS node, the outgoing interface and the next hop. There are only IS-IS nodes at IS-IS level 2. The output of this command shows that data traffic from PE-1 to PE-2 uses interface "int-PE-1-PE-2" to PE-2, while, for example, data traffic from PE-1 to P-3 uses interface "int-PE-1-P-3" to P-3.

```

*A:PE-1# show router isis 0 topology

=====
Rtr Base ISIS Instance 0 Topology Table
=====
Node                               Interface                               Nexthop
-----
IS-IS IP paths (MT-ID 0),  Level 2
-----
PE-2.00                            int-PE-1-PE-2                        PE-2
P-3.00                             int-PE-1-P-3                         P-3
P-4.00                             int-PE-1-P-4                         P-4
=====

```

The **show router isis 0 topology detail** command lists the IS-IS nodes, and for each IS-IS node, the next hop, the outgoing interface, and the metric (in this case, IS-IS level 2) that applies.

```

*A:PE-1# show router isis 0 topology detail

=====
Rtr Base ISIS Instance 0 Topology Table
=====
IS-IS IP paths (MT-ID 0),  Level 2
-----
Node      : PE-2.00
Nexthop   : PE-2
Interface : int-PE-1-PE-2
SNPA      : none                               Metric    : 8

Node      : P-3.00
Nexthop   : P-3
Interface : int-PE-1-P-3
SNPA      : none                               Metric    : 8

Node      : P-4.00
Nexthop   : P-4
Interface : int-PE-1-P-4
SNPA      : none                               Metric    : 15
=====

```



Verify the IS-IS data base on PE-1 with **show router isis 0 database detail**. The output of this command (shortened here for PE-1 and PE-2, and omitted for P-3 and P-4) provides information about each IS-IS-enabled router. For each uniquely identified IS-IS-enabled router, the SRv6 information indicates:

- the IS-IS-advertised router capabilities
- the IS-IS topology details
- the IPv4 and IPv6 reachability details
- the advertised SRv6 locator TLV
- the advertised configured SRv6 End SID

There is only IS-IS information at IS-IS level 2. On each IS-IS interface, the IS-IS level 2 metrics for IPv4 and IPv6 are identical.

Only the default metric-based SPF algorithm instance 0 is in use, as listed in the SR Alg sub-TLV of the Router Cap TLVs.

On PE-1 and PE-2, the SRv6 locator prefix for algorithm 0, with their End SRv6 SID are present.

```
*A:PE-1# show router isis 0 database detail

=====
Rtr Base ISIS Instance 0 Database (detail)
=====
---snip---
Displaying Level 2 database
-----
LSP ID      : PE-1.00-00                                Level      : L2
---snip---
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 : 3 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
  IPv6 Address : 2001:db8::168:14:1
TE IS Nbrs :
Nbr      : PE-2.00
Default Metric : 8
Sub TLV Len : 36
IPv6 Addr  : 2001:db8::168:12:1
Nbr IPv6   : 2001:db8::168:12:2
TE IS Nbrs :
Nbr      : P-3.00
Default Metric : 8
Sub TLV Len : 36
IPv6 Addr  : 2001:db8::168:13:1
```

```
Nbr IPv6 : 2001:db8::168:13:2
TE IS Nbrs :
  Nbr : P-4.00
  Default Metric : 15
  Sub TLV Len : 36
  IPv6 Addr : 2001:db8::168:14:1
  Nbr IPv6 : 2001:db8::168:14:2
IPv6 Reach:
  Metric: ( I ) 0
  Prefix : 2001:db8::2:1/128
  Metric: ( I ) 8
  Prefix : 2001:db8::168:12:0/126
  Metric: ( I ) 8
  Prefix : 2001:db8::168:13:0/126
  Metric: ( I ) 15
  Prefix : 2001:db8::168:14:0/126
  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
---snip---
TLVs :
Area Addresses:
  Area Address : (3) 49.0001
Supp Protocols:
  Protocols : IPv4
  Protocols : IPv6
IS-Hostname : PE-2
Router ID :
  Router ID : 2.2.2.2
TE Router ID v6 :
  Router ID : 2001:db8::2:2
Router Cap : 2.2.2.2, 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 : 3 SRH-MAX-END-
D : 9
---snip---
TE IS Nbrs :
  Nbr : PE-1.00
  Default Metric : 9
---snip---
TE IS Nbrs :
  Nbr : P-3.00
  Default Metric : 17
---snip---
TE IS Nbrs :
  Nbr : P-4.00
  Default Metric : 9
---snip---
---snip---
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
---snip---
Level (2) LSP Count : 4
-----
---snip---
SABM-flags Flags:  R = RSVP-TE
                   S = SR-TE
                   F = LFA
                   X = FLEX-ALGO
FAD-flags Flags:   M = Prefix Metric
=====

```

PE-2 advertises to PE-1 the information for network prefix 172.16.222.1/32, as listed in the RIB In Entries section in the following example. PE-1 acts in a similar way as PE-2 for network prefix 172.16.211.1/32, as listed in the RIB Out Entries section.

The following output shows the corresponding VPN-IPv4 BGP routes on PE-1:

```

*A:PE-1# show router bgp routes vpn-ipv4 hunt
=====
BGP Router ID:1.1.1.10      AS:64500      Local AS:64500
=====
---snip---
=====
BGP VPN-IPv4 Routes
=====
-----
RIB In Entries
-----
Network      : 172.16.222.1/32
Nextthop    : 2001:db8::2:2
Route Dist.    : 192.0.2.2:6      VPN Label    : 524288
Path Id       : None
From        : 2001:db8::2:2
Res. Nextthop : n/a
Local Pref.   : 100
Aggregator AS : None              Interface Name : int-PE-1-PE-2
Atomic Aggr. : Not Atomic        Aggregator    : None
AIGP Metric   : None             MED           : None
Connector     : None             IGP Cost    : 8
Community    : target:64506:6
Cluster      : No Cluster Members
Originator Id : None             Peer Router Id : 2.2.2.10
Fwd Class    : None             Priority      : None
Flags        : Used Valid Best IGP
---snip---
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)
---snip---
VPRN Imported  : 6
-----
RIB Out Entries
-----
Network      : 172.16.211.1/32
Nextthop    : 2001:db8::2:1
Route Dist.    : 192.0.2.1:6      VPN Label    : 524288
Path Id       : None
To         : 2001:db8::2:2

```

```

Res. Nexthop      : n/a
Local Pref.      : 100
Aggregator AS    : None
Atomic Aggr.     : Not Atomic
AIGP Metric      : None
Connector        : None
Community        : target:64506:6
Cluster          : No Cluster Members
Originator Id    : None
Origin           : IGP
Peer Router Id   : 2.2.2.10
---snip---
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)
---snip---
-----
Routes : 2
=====

```

Verify the IPv6 route table on PE-1. The IPv6 route table has routes to the local and remote SRv6 locators and to the local SRv6 End function SID. The SRv6 locator prefix of PE-2 is reached via an SRv6 tunnel using IS-IS. The routes with protocol **"SRV6"** correspond with the locally configured SRv6 locator prefix of PE-1 and the locally configured SRv6 End function.

```

*A: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   00h11m13s  3
    fe80::201- "_tmnx_fpe_2.a"         0
2001:db8:aaaa:101:0:1000::/128 Local  SRV6   00h11m13s  3
    Black Hole                           0
2001:db8:aaaa:102::/64      Remote ISIS   00h10m26s 18
    2001:db8:aaaa:102::/64 (tunneled:SRV6-ISIS) 8
-----
No. of Routes: 13
---snip---
=====

```

Verify that the tunnel from PE-1 to the SRv6 locator prefix of PE-2 is an SRv6 tunnel that uses the "int-PE-1-PE-2" interface. A similar verification can be performed for the other direction. Interface "int-PE-1-PE-2" is configured on port 1/1/c1/1:1000.

```

*A:PE-1# show router fp-tunnel-table 1 ipv6

=====
IPv6 Tunnel Table Display
---snip---
=====
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 also that IPv6 data traffic is possible between the local VPRN on PE-1 and the remote VPRN on PE-2:

```

*A:PE-1# ping router 6 2001:db8:222::1
PING 2001:db8:222::1 56 data bytes
64 bytes from 2001:db8:222::1 icmp_seq=1 hlim=64 time=1.84ms.
---snip---
---- 2001:db8:222::1 PING Statistics ----
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min = 1.54ms, avg = 1.72ms, max = 1.84ms, stddev = 0.104ms

*A:PE-1# traceroute router 6 2001:db8:222::1
traceroute to 2001:db8:222::1, 30 hops max, 60 byte packets
 1 2001:db8:222::1 (2001:db8:222::1)  1.68 ms  1.84 ms  1.65 ms

```

For IPv6 data traffic, the VPRN routing table shows the next hop and the applicable IGP metric for the route to 2001:db8:222::1.

```

*A:PE-1# show router 6 route-table 2001:db8:222::1
=====
IPv6 Route Table (Service: 6)
=====
Dest Prefix[Flags]                Type   Proto   Age      Pref
  Next Hop[Interface Name]                Metric
-----
2001:db8:222::1/128              Remote BGP VPN 00h01m52s 170
  2001:db8:aaaa:102:7fff:f000:: (tunneled:SRV6)      8
-----
No. of Routes: 1
---snip---
=====

```

The next hop is the End.DT6 SRv6 SID of the SRv6 locator "PE-2\_loc" for the VPRN on PE-2, which PE-1 learns from a BGP update from PE-2. The SRv6 tunnel to this next hop has label 524287.

The applicable IGP metric is also 8, which again corresponds with the IS-IS level 2 "int-PE-1-PE-2" interface metric value.

PE-2 advertises to PE-1 the information for network prefix 2001:db8:222::1/128, as listed in the RIB In Entries section in the following example. PE-1 acts in a similar way as PE-2 for network prefix 2001:db8:211::1/128, as listed in the RIB Out Entries section.

The following output shows the corresponding VPN-IPv6 BGP routes on PE-1:

```

*A:PE-1# show router bgp routes vpn-ipv6 hunt
=====
BGP Router ID:1.1.1.10          AS:64500      Local AS:64500
=====
---snip---
=====

```

**BGP VPN-IPv6 Routes**

**RIB In Entries**

```

Network      : 2001:db8:222::1/128
Nextthop    : 2001:db8::2:2
Route Dist. : 192.0.2.2:6          VPN Label   : 524287
Path Id     : None
From        : 2001:db8::2:2
Res. Nextthop : 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     : 8
Connector   : None
Community   : target:64506:6
Cluster     : No Cluster Members
Originator Id : None              Peer Router Id : 2.2.2.10
Fwd Class   : None                Priority      : None
Flags       : Used Valid Best IGP
---snip---
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)
---snip---
VPRN Imported : 6
    
```

**RIB Out Entries**

```

Network      : 2001:db8:211::1/128
Nextthop    : 2001:db8::2:1
Route Dist. : 192.0.2.1:6          VPN Label   : 524287
Path Id     : None
To          : 2001:db8::2:2
Res. Nextthop : 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:64506:6
Cluster     : No Cluster Members
Originator Id : None              Peer Router Id : 2.2.2.10
Origin      : IGP
---snip---
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)
---snip---
    
```

Routes : 2

## Configure a delay-based flexible algorithm

Define a Flex-algorithm definition "FAD\_delay" that takes delay as its metric. The FAD can reside on any IS-IS-enabled router. In this example, it resides on PE-1.

```
*A:PE-1# configure router Base
    flexible-algorithm-definitions
    flex-algo "FAD_delay" create
        description "FAD_delay_based"
        metric-type delay
        no shutdown
    exit
exit all
```

Configure PE-1 so as to advertise the FAD "FAD\_delay" in Flex-Algorithm instance 128 (possible values between 128 and 255) and to participate in the Flex-Algorithm instance 128.

```
*A:PE-1# configure router Base isis 0
    flexible-algorithms
    flex-algo 128
        advertise "FAD_delay"
        participate
    exit
exit
no shutdown
exit all
```

Ensure that the other IS-IS-enabled routers that must support the Flex-Algorithm instance participate in this Flex-Algorithm instance.

```
*A:PE-2#/*A:P-3#/*A:P-4# configure router Base isis 0
    flexible-algorithms
    flex-algo 128
        participate
    exit
exit
no shutdown
exit all
```

Define a new and unique SRv6 locator prefix to this new Flex-Algorithm instance. A separate SRv6 locator is needed for each Flex-Algorithm instance. So, for the Flex-Algorithm instance 128, configure SRv6 locator "PE-2\_loc\_FAD128" on PE-2. Perform a similar configuration for SRv6 locator "PE-1\_loc\_FAD128" with ip-prefix 2001:db8:a128:101::/64 on PE-1.

```
*A:PE-2# configure router Base segment-routing segment-routing-v6
    locator "PE-2_loc_FAD128"
        block-length 48
        function-length 20
        algorithm 128
        prefix
            ip-prefix 2001:db8:a128:102::/64
    exit
no shutdown
exit all
```

For SRv6 locator "PE-2\_loc\_FAD128" on PE-2, use FPE 2 as the SRv6 termination FPE in the **router Base segment-routing segment-routing-v6 locator <locator-name>** context and configure the SRv6

End function (equivalent to an IPv4 node SID) in the **router Base segment-routing segment-routing-v6 base-routing-instance locator <locator-name>** context. Perform a similar configuration on PE-1, for SRv6 locator "PE-1\_loc\_FAD128".

```
*A:PE-2# configure router Base segment-routing
segment-routing-v6
  locator "PE-2_loc_FAD128"
  termination-fpe 2
  no shutdown
exit
base-routing-instance
  locator "PE-2_loc_FAD128"
  function
  end 1
  srh-mode usp
exit
exit
exit
exit all
```

In the **router Base isis 0 segment-routing-v6 locator <locator-name>** context on PE-2, configure the IS-IS level capability for SRv6 locator "PE-2\_loc\_FAD128" and enable SRv6 in the IS-IS context. Perform a similar configuration on PE-1, for SRv6 locator "PE-1\_loc\_FAD128".

```
*A:PE-2# configure router Base isis 0
segment-routing-v6
  locator "PE-2_loc_FAD128"
  level-capability level-2
  level 2
exit
exit
no shutdown
exit all
```

The **show router isis 0 segment-routing-v6 locator** command lists the local and remote SRv6 locator prefixes for all applied algorithms (0 and 128).

```
*A:PE-2# show router isis 0 segment-routing-v6 locator

=====
Rtr Base ISIS Instance 0 SRv6 Locator Table
=====
Prefix                               AdvRtr      MT      Lvl/Typ
AttributeFlags                       Tag         Flags   Algo
-----
2001:db8:a128:101::/64                PE-1        0       2/Int.
-                                     0           -       128
2001:db8:a128:102::/64                PE-2        0       2/Int.
-                                     0           -       128
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: 4
-----
--- snip ---
=====
```



On PE-2, in the **service vprn segment-routing-v6 locator** context, configure End.DT4 and End.DT6 functions for SRv6 locator "*PE-2\_loc\_FAD128*".

On PE-2, in the **service vprn <service-id> bgp-ipvpn segment-routing-v6** context, use the SRv6 locator "*PE-2\_loc\_FAD128*" as the default locator, instead of the earlier SRv6 locator "*PE-2\_loc*". The BGP IPVPN SRv6 instance for the VPRN must be shut down to allow this replacement. Perform a similar configuration on PE-1, for SRv6 locator "*PE-1\_loc\_FAD128*".

```
*A:PE-2# configure service
      vprn 6
        segment-routing-v6 1
          locator "PE-2_loc_FAD128"
            function
              end-dt4
              end-dt6
            exit
          exit
        exit
      bgp-ipvpn
        segment-routing-v6
          shutdown
          srv6-instance 1 default-locator "PE-2_loc_FAD128"
          no shutdown
        exit
      exit
    no shutdown
  exit all
```

## Verify data traffic

At this point, using **ping** and **traceroute** commands, verify that data traffic between the local VPRN on PE-1 and the remote VPRN on PE-2 uses the Flex-Algorithm.

For IPv4 data traffic:

```
*A:PE-1# ping router 6 172.16.222.1
PING 172.16.222.1 56 data bytes
64 bytes from 172.16.222.1: icmp_seq=1 ttl=64 time=2.47ms.
---snip---
---- 172.16.222.1 PING Statistics ----
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min = 1.93ms, avg = 2.35ms, max = 2.51ms, stddev = 0.210ms

*A:PE-1# traceroute router 6 172.16.222.1
traceroute to 172.16.222.1, 30 hops max, 40 byte packets
 1 172.16.222.1 (172.16.222.1)  2.79 ms  2.43 ms  2.41 ms
```

For IPv6 data traffic:

```
*A:PE-1# ping router 6 2001:db8:222::1
PING 2001:db8:222::1 56 data bytes
64 bytes from 2001:db8:222::1 icmp_seq=1 hlim=64 time=2.31ms.
---snip---
---- 2001:db8:222::1 PING Statistics ----
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min = 2.31ms, avg = 2.50ms, max = 2.71ms, stddev = 0.150ms

*A:PE-1# traceroute router 6 2001:db8:222::1
traceroute to 2001:db8:222::1, 30 hops max, 60 byte packets
```

```
1 2001:db8:222::1 (2001:db8:222::1) 2.50 ms 2.48 ms 2.40 ms
```

This data traffic uses the SRv6 tunnels over the links between PE-1 and P-3, P-3 and P-4, and P-4 and PE-2.

For IPv4 data traffic, the VPRN routing table shows the next hop and the applicable metric for the route to 172.16.222.1. In this example, the End.DT4 SID is copied into the IPv6 DA field of the tunneled packet.

```
*A:PE-1# show router 6 route-table 172.16.222.1
=====
Route Table (Service: 6)
=====
Dest Prefix[Flags]                               Type   Proto   Age           Pref
  Next Hop[Interface Name]                       Metric
-----
172.16.222.1/32                                  Remote BGP VPN 00h02m07s 170
          2001:db8:a128:102:7fff:e000:: (tunneled:SRV6)          111
-----
No. of Routes: 1
---snip---
```

The tunnel next hop is the End.DT4 SRv6 SID of the SRv6 **locator** "PE-2\_loc\_FAD128" for the VPRN on PE-2, which PE-1 learns from a BGP update from PE-2. The SRv6 tunnel to this next hop has label 524286, which is the transposed SRv6 End.DT4 SID function value 0x7fffe.

```
*A:PE-2# show router segment-routing-v6 local-sid
=====
Segment Routing v6 Local SIDs
=====
SID                               Type           Function
Locator
Context
-----
2001:db8:a128:102:0:1000::        End             1
  PE-2_loc_FAD128
  Base
2001:db8:a128:102:7fff:d000::    End.DT6         524285
  PE-2_loc_FAD128
  SvcId: 6 Name: VPRN_6
2001:db8:a128:102:7fff:e000::    End.DT4         524286
  PE-2_loc_FAD128
  SvcId: 6 Name: VPRN_6
---snip---
```

The **show router isis 0 topology flex-algo 128** command lists the IS-IS nodes in the topology, and for each IS-IS node, the outgoing interface and the next hop. There are only IS-IS nodes at IS-IS level 2. The output of this command shows that data traffic from PE-1 to all IS-IS-enabled routers that participate in the Flex-Algorithm instance 128 (PE-2, P-3, and P-4) uses interface "int-PE-1-P-3" to P-3.

```
*A:PE-1# show router isis 0 topology flex-algo 128
=====
Rtr Base ISIS Instance 0 Flex-Algo 128 Topology Table
```

```

=====
Node                               Interface                               Nexthop
-----
IS-IS IP paths (MT-ID 0),  Level 2
-----
PE-2.00                            int-PE-1-P-3                            P-3
P-3.00                              int-PE-1-P-3                            P-3
P-4.00                              int-PE-1-P-3                            P-3
=====

```

The applicable metric is 111, which corresponds with the sum of the static link delays that are configured on the router interfaces "int-PE-1-P-3", "int-P-3-P-4", and "int-P-4-PE-2".

The **show router isis 0 topology flex-algo 128 detail** command lists the IS-IS nodes in the topology, and for each IS-IS node, the next hop, the outgoing interface, and the metric (in this case, link delay) that applies.

```

*A:PE-1# show router isis 0 topology flex-algo 128 detail

=====
Rtr Base ISIS Instance 0 Flex-Algo 128 Topology Table
=====
IS-IS IP paths (MT-ID 0),  Level 2
-----
Node      : PE-2.00
Nexthop   : P-3
Interface : int-PE-1-P-3
SNPA      : none
Metric    : 111

Node      : P-3.00
Nexthop   : P-3
Interface : int-PE-1-P-3
SNPA      : none
Metric    : 1

Node      : P-4.00
Nexthop   : P-3
Interface : int-PE-1-P-3
SNPA      : none
Metric    : 11

=====

*A:P-3# show router isis 0 topology flex-algo 128 detail

=====
Rtr Base ISIS Instance 0 Flex-Algo 128 Topology Table
=====
IS-IS IP paths (MT-ID 0),  Level 2
-----
---snip---
Node      : PE-2.00
Nexthop   : P-4
Interface : int-P-3-P-4
SNPA      : none
Metric    : 110

Node      : P-4.00
Nexthop   : P-4
Interface : int-P-3-P-4
SNPA      : none
Metric    : 10

=====

```

```
*A:P-4# show router isis 0 topology flex-algo 128 detail

=====
Rtr Base ISIS Instance 0 Flex-Algo 128 Topology Table
=====
-----
IS-IS IP paths (MT-ID 0),   Level 2
-----
---snip---
Node       : PE-2.00
Nexthop    : PE-2
Interface  : int-P-4-PE-2
SNPA       : none                               Metric    : 100
---snip---
=====
```

The IS-IS database on PE-1 contains more information, which relates to the use of the Flex-Algorithm.

There are additional link delay metrics (identical for IPv4 and IPv6) for each IS-IS-enabled router, as listed in the TE APP LINK ATTR sub-TLVs. The non-legacy Standard Application Bit Mask (SABM) flag value X indicates that they are associated with the Flex-Algorithm. Only the End SRv6 SIDs of the SRv6 locators are present.

Next to the default metric-based SPF, the Flex-Algorithm instance 128 is also in use, as listed in the SR Alg sub-TLV of the Router Cap TLVs.

The FAD sub-TLV of the PE-1 Router Cap TLV contains the delay-based definition for the Flex-Algorithm instance 128, which only router PE-1 advertises. The FAD flag value M indicates that the delay is a prefix metric.

On PE-1 and PE-2, next to the base SRv6 locator (for algorithm 0), with its End SRv6 SID, there is an additional SRv6 locator for the Flex-Algorithm instance 128, with its End SRv6 SID. This additional SRv6 locator indicates the prefix.

The **show router isis 0 database detail** command output on PE-1 is shown below, with separate entries for each IS-IS-enabled router.

For PE-1:

```
*A:PE-1# show router isis 0 database detail

=====
Rtr Base ISIS Instance 0 Database (detail)
=====
---snip---
Displaying Level 2 database
-----
LSP ID      : PE-1.00-00                               Level      : L2
---snip---
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, 128
Node MSD Cap: BMI : 0 SRH-MAX-SL : 10 SRH-MAX-END-POP : 9 SRH-MAX-H-ENCAPS : 3 SRH-MAX-END-
D : 9
  FAD Sub-Tlv:
    Flex-Algorithm   : 128
    Metric-Type     : delay
    Calculation-Type : 0
    Priority         : 100
    Flags: M
I/F Addresses IPv6 :
  IPv6 Address      : 2001:db8::2:1
  IPv6 Address      : 2001:db8::168:12:1
  IPv6 Address      : 2001:db8::168:13:1
  IPv6 Address      : 2001:db8::168:14:1
TE IS Nbrs       :
  Nbr               : PE-2.00
  Default Metric    : 8
  Sub TLV Len       : 36
  IPv6 Addr         : 2001:db8::168:12:1
  Nbr IPv6          : 2001:db8::168:12:2
TE IS Nbrs       :
  Nbr               : P-3.00
  Default Metric    : 8
  Sub TLV Len       : 51
  IPv6 Addr         : 2001:db8::168:13:1
  Nbr IPv6          : 2001:db8::168:13:2
  TE APP LINK ATTR :
    SABML-flag:Non-Legacy SABM-flags: X
    Delay Min : 1 Max : 1
TE IS Nbrs       :
  Nbr               : P-4.00
  Default Metric    : 15
  Sub TLV Len       : 36
  IPv6 Addr         : 2001:db8::168:14:1
  Nbr IPv6          : 2001:db8::168:14:2
IPv6 Reach:
  Metric: ( I ) 0
  Prefix   : 2001:db8::2:1/128
  Metric: ( I ) 8
  Prefix   : 2001:db8::168:12:0/126
  Metric: ( I ) 8
  Prefix   : 2001:db8::168:13:0/126
  Metric: ( I ) 15
  Prefix   : 2001:db8::168:14:0/126
  Metric: ( I ) 0
  Prefix   : 2001:db8:aaaa:101::/64
SRv6 Locator   :
  MT ID : 0
  Metric: ( ) 0 Algo:128
  Prefix : 2001:db8:a128:101::/64
  Sub TLV :
    End-SID : 2001:db8:a128:101:0:1000::, flags:0x0, endpoint:End-USP
  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
---snip---
Level (2) LSP Count : 4
-----
---snip---
SABM-flags Flags:  R = RSVP-TE
                   S = SR-TE
                   F = LFA
```

```
          X = FLEX-ALGO
FAD-flags Flags:  M = Prefix Metric
=====
```

For PE-2:

```
*A:PE-1# show router isis 0 database detail

=====
Rtr Base ISIS Instance 0 Database (detail)
=====
---snip---
Displaying Level 2 database
-----
---snip---
LSP ID      : PE-2.00-00                Level      : L2
---snip---
TLVs :
---snip---
Router Cap : 2.2.2.2, D:0, S:0
TE Node Cap : B E M P
SRv6 Cap: 0x0000
SR Alg: metric based SPF, 128
Node MSD Cap: BMI : 0 SRH-MAX-SL : 10 SRH-MAX-END-POP : 9 SRH-MAX-H-ENCAPS : 3 SRH-MAX-END-
D : 9
---snip---
TE IS Nbrs :
  Nbr      : PE-1.00
  Default Metric : 9
---snip---
TE IS Nbrs :
  Nbr      : P-3.00
  Default Metric : 17
---snip---
TE IS Nbrs :
  Nbr      : P-4.00
  Default Metric : 9
---snip---
TE APP LINK ATTR      :
  SABML-flag:Non-Legacy SABM-flags: X
  Delay Min : 200 Max : 200
---snip---
SRv6 Locator :
  MT ID : 0
  Metric: ( ) 0 Algo:128
  Prefix : 2001:db8:a128:102::/64
  Sub TLV :
    End-SID : 2001:db8:a128:102:0:1000::, flags:0x0, endpoint:End-USP
  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
---snip---
=====
```

For P-3:

```
*A:PE-1# show router isis 0 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---
Router Cap : 3.3.3.3, D:0, S:0
  TE Node Cap : B E M P
  SRv6 Cap: 0x0000
  SR Alg: metric based SPF, 128
  Node MSD Cap: BMI : 0 SRH-MAX-SL : 10 SRH-MAX-END-POP : 9 SRH-MAX-H-ENCAPS : 3 SRH-MAX-END-
D : 9
---snip---
TE IS Nbrs :
  Nbr : PE-1.00
  Default Metric : 10
---snip---
TE APP LINK ATTR      :
  SABML-flag:Non-Legacy SABM-flags: X
  Delay Min : 2 Max : 2
TE IS Nbrs :
  Nbr : PE-2.00
  Default Metric : 18
---snip---
TE IS Nbrs :
  Nbr : P-4.00
  Default Metric : 10
---snip---
TE APP LINK ATTR      :
  SABML-flag:Non-Legacy SABM-flags: X
  Delay Min : 10 Max : 10
---snip---
=====

```

For P-4:

```

*A:PE-1# show router isis 0 database detail

=====
Rtr Base ISIS Instance 0 Database (detail)
=====
---snip---
Displaying Level 2 database
-----
---snip---
LSP ID      : P-4.00-00                                Level      : L2
---snip---
TLVs :
---snip---
Router Cap : 4.4.4.4, D:0, S:0
  TE Node Cap : B E M P
  SRv6 Cap: 0x0000
  SR Alg: metric based SPF, 128
  Node MSD Cap: BMI : 0 SRH-MAX-SL : 10 SRH-MAX-END-POP : 9 SRH-MAX-H-ENCAPS : 3 SRH-MAX-END-
D : 9
---snip---
TE IS Nbrs :
  Nbr : PE-1.00
  Default Metric : 16
---snip---
TE IS Nbrs :
  Nbr : PE-2.00

```

```

Default Metric : 9
---snip---
TE APP LINK ATTR :
  SABML-flag:Non-Legacy SABM-flags: X
  Delay Min : 100 Max : 100
TE IS Nbrs :
  Nbr : P-3.00
  Default Metric : 9
---snip---
TE APP LINK ATTR :
  SABML-flag:Non-Legacy SABM-flags: X
  Delay Min : 20 Max : 20
---snip---
=====

```

PE-2 advertises to PE-1 payload prefix 172.16.222.1/32, as listed in the RIB In Entries section in the following example. PE-1 computes the applicable metric 111, which corresponds with the sum of the link delays that are configured on the router interfaces "int-PE-1-P-3", "int-P-3-P-4", and "int-P-4-PE-2". PE-1 advertises to PE-2 payload prefix 172.16.211.1/32, as listed in the RIB Out Entries section in the following example. PE-2 computes the applicable metric 222, which corresponds with the sum of the link delays that are configured on the router interfaces "int-PE-2-P-4", "int-P-4-P-3", and "int-P-3-PE-1".

The following output shows the corresponding VPN-IPv4 BGP routes on PE-1:

```

*A:PE-1# show router bgp routes vpn-ipv4 hunt
=====
BGP Router ID:1.1.1.10      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.  : 192.0.2.2:6      VPN Label    : 524286
Path Id      : None
From         : 2001:db8::2:2
Res. NextHop : n/a
Local Pref.  : 100
Aggregator AS : None      Interface Name : int-PE-1-PE-2
Atomic Aggr. : Not Atomic  Aggregator     : None
AIGP Metric  : None      MED            : None
Connector    : None      IGP Cost       : 111
Community    : target:64506:6
Cluster      : No Cluster Members
Originator Id : None      Peer Router Id : 2.2.2.10
Fwd Class    : None      Priority        : None
Flags        : Used Valid Best IGP
---snip---
SRv6 TLV Type : SRv6 L3 Service TLV (5)
SRv6 SubTLV   : SRv6 SID Information (1)
Sid           : 2001:db8:a128:102::
Full Sid      : 2001:db8:a128:102:7fff:e000::
Behavior      : End.DT4 (19)
SRv6 SubSubTLV : SRv6 SID Structure (1)
---snip---
VPRN Imported : 6
-----

```



```

RIB Out Entries
-----
Network       : 172.16.211.1/32
Nexthop       : 2001:db8::2:1
Route Dist.   : 192.0.2.1:6          VPN Label    : 524286
Path Id       : None
To            : 2001:db8::2:2
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:64506:6
Cluster       : No Cluster Members
Originator Id : None                 Peer Router Id : 2.2.2.10
Origin        : IGP
---snip---
SRv6 TLV Type : SRv6 L3 Service TLV (5)
SRv6 SubTLV   : SRv6 SID Information (1)
Sid           : 2001:db8:a128:101::
Full Sid      : 2001:db8:a128:101:7fff:e000::
Behavior      : End.DT4 (19)
SRv6 SubSubTLV : SRv6 SID Structure (1)
---snip---
-----
Routes : 2
=====

```

Verify the IPv6 route table on PE-1. The IPv6 route table has an additional route to the learned remote SRv6 locator for the Flex-Algorithm instance 128. This remotely configured SRv6 locator prefix of PE-2 is reached via an SRv6 tunnel.

```

*A: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:a128:101::/64             Local  SRV6    00h02m30s    3
    fe80::201- "_tmnx_fpe_2.a"    0
2001:db8:a128:101:0:1000::/128    Local  SRV6    00h02m30s    3
    Black Hole                      0
2001:db8:a128:102::/64           Remote ISIS    00h01m55s    18
    2001:db8:a128:102::/64 (tunneled:SRV6-ISIS)
111
2001:db8:aaaa:101::/64             Local  SRV6    00h18m00s    3
    fe80::201- "_tmnx_fpe_2.a"    0
2001:db8:aaaa:101:0:1000::/128    Local  SRV6    00h18m00s    3
    Black Hole                      0
2001:db8:aaaa:102::/64             Remote  ISIS    00h17m13s    18
    2001:db8:aaaa:102::/64 (tunneled:SRV6-ISIS)
    8
-----
No. of Routes: 16
---snip---
=====

```

Verify that the tunnel from PE-1 to the remote locator is an SRv6 tunnel that uses the "int-PE-1-P-3" interface. Perform a similar verification for the tunnel from PE-2, where the SRv6 tunnel to the remote locator uses the "int-PE-2-P-4" interface. Interface "int-PE-1-P-3" is configured on port 1/1/c2/1:1000.

```
*A:PE-1# show router fp-tunnel-table 1 ipv6
=====
IPv6 Tunnel Table Display
---snip---
=====
Destination                                Protocol      Tunnel-ID
Lbl/SID
  NextHop                                  Intf/Tunnel
Lbl/SID (backup)
  NextHop (backup)
-----
2001:db8:a128:102::/64                      SRV6         524290
-
  fe80::612:1ff:fe01:b-"int-PE-1-P-3"      1/1/c2/1:1000
2001:db8:aaa:102::/64                      SRV6         524289
-
  fe80::60e:1ff:fe01:1-"int-PE-1-PE-2"    1/1/c1/1:1000
-----
Total Entries : 2
=====
```

For IPv6 data traffic, the VPRN routing table shows the next hop and the applicable metric for the route to 2001:db8:222::1.

```
*A:PE-1# show router 6 route-table 2001:db8:222::1
=====
IPv6 Route Table (Service: 6)
=====
Dest Prefix[Flags]                          Type      Proto      Age          Pref
Next Hop[Interface Name]                    Metric
-----
2001:db8:222::1/128                          Remote    BGP VPN    00h02m07s   170
          2001:db8:a128:102:7fff:d000:: (tunneled:SRV6)
                                          111
-----
No. of Routes: 1
---snip---
=====
```

The next hop is the End.DT6 SRv6 SID of the SRv6 **locator** "PE-2\_loc\_FAD128" for the VPRN on PE-2, which PE-1 learns from a BGP update from PE-2. The SRv6 tunnel to this next hop has label 524285.

The applicable metric is also 111, which again corresponds with the sum of the link delays that are configured on the router interfaces "int-PE-1-P-3", "int-P-3-P-4", and "int-P-4-PE-2".

PE-2 advertises to PE-1 the information for network prefix 2001:db8:222::1/128, as listed in the RIB In Entries section in the following example. PE-1 computes the applicable metric 111, which corresponds with the sum of the link delays that are configured on the router interfaces "int-PE-1-P-3", "int-P-3-P-4", and "int-P-4-PE-2". PE-1 advertises to PE-2 payload prefix 2001:db8:211::1/128, as listed in the RIB Out Entries section in the following example. PE-2 computes the applicable metric 222, which corresponds with the sum of the link delays that are configured on the router interfaces "int-PE-2-P-4", "int-P-4-P-3", and "int-P-3-PE-1".

The following output shows the corresponding VPN-IPv6 BGP routes on PE-1:

```
*A:PE-1# show router bgp routes vpn-ipv6 hunt
=====
BGP Router ID:1.1.1.10      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.  : 192.0.2.2:6      VPN Label    : 524285
Path Id      : None
From         : 2001:db8::2:2
Res. NextHop : n/a
Local Pref.  : 100
Aggregator AS : None           Interface Name : int-PE-1-PE-2
Atomic Aggr. : Not Atomic   Aggregator    : None
AIGP Metric  : None         MED           : None
Connector    : None         IGP Cost      : 111
Community    : target:64506:6
Cluster      : No Cluster Members
Originator Id : None           Peer Router Id : 2.2.2.10
Fwd Class    : None         Priority       : None
Flags        : Used Valid Best IGP
---snip---
SRv6 TLV Type : SRv6 L3 Service TLV (5)
SRv6 SubTLV   : SRv6 SID Information (1)
Sid           : 2001:db8:a128:102::
Full Sid      : 2001:db8:a128:102:7fff:d000::
Behavior      : End.DT6 (18)
SRv6 SubSubTLV : SRv6 SID Structure (1)
---snip---
VPRN Imported : 6
-----
RIB Out Entries
-----
Network      : 2001:db8:211::1/128
NextHop      : 2001:db8::2:1
Route Dist.  : 192.0.2.1:6      VPN Label    : 524285
Path Id      : None
To           : 2001:db8::2:2
Res. NextHop : n/a
Local Pref.  : 100
Aggregator AS : None           Interface Name : NotAvailable
Atomic Aggr. : Not Atomic   Aggregator    : None
AIGP Metric  : None         MED           : None
Connector    : None         IGP Cost      : n/a
Community    : target:64506:6
Cluster      : No Cluster Members
Originator Id : None           Peer Router Id : 2.2.2.10
Origin       : IGP
---snip---
SRv6 TLV Type : SRv6 L3 Service TLV (5)
SRv6 SubTLV   : SRv6 SID Information (1)
Sid           : 2001:db8:a128:101::
Full Sid      : 2001:db8:a128:101:7fff:d000::
Behavior      : End.DT6 (18)
SRv6 SubSubTLV : SRv6 SID Structure (1)
```

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

## Configure TE metrics on the router interfaces

For example for PE-1. A similar configuration applies for the other routers.

```
*A:PE-1# configure
router Base
 mpls
   interface "int-PE-1-PE-2"
     te-metric 1500
     no shutdown
   exit
   interface "int-PE-1-P-3"
     te-metric 10
     no shutdown
   exit
   interface "int-PE-1-P-4"
     te-metric 1600
     no shutdown
   exit
 no shutdown
 exit
 rsvp
 no shutdown
 exit
 exit all
```

## Configure a TE-metric-based flexible algorithm

Define a Flex-Algorithm definition "FAD\_te\_metric" that uses TE metric metric type. The FAD can reside on any IS-IS-enabled router. In this example, it resides on PE-1.

```
*A:PE-1# configure router Base flexible-algorithm-definitions # strictly needed on only 1
router in ISIS level
 flex-algo "FAD_te_metric" create
   description "FAD_te_metric_based"
   metric-type te-metric
 exit all
```

Configure PE-1 so as to advertise the FAD "FAD\_te\_metric" in Flex-Algorithm instance 128 and to participate in the Flex-Algorithm instance 128.



**Note:** SR OS 22.10 supports multiple concurrent Flex-Algorithm instances, with different instance values between 128 and 255. While "FAD\_te\_metric" would be associated with a new Flex-Algorithm instance 129 in a deployment scenario, "FAD\_te\_metric" in the example setup replaces "FAD\_delay" for Flex-Algorithm instance 128, so that, for the sake of brevity, the SRv6 locators and labels that were in use earlier for the delay-based flexible algorithm scenario remain in use for the TE-metric-based flexible algorithm scenario.

```
*A:PE-1# configure router Base isis 0
 flexible-algorithms
```

```

flex-algo 128
  advertise "FAD_te_metric"
  participate
  exit
exit
no shutdown
exit all

```

Ensure that the other IS-IS-enabled routers that must support the Flex-Algorithm instance participate in this Flex-Algorithm instance.

```

*A:PE-2#/*A:P-3#/*A:P-4# configure router Base isis 0
  flexible-algorithms
  flex-algo 128
  participate
  exit
exit
no shutdown
exit all

```

## Verify data traffic

At this point, using **ping** and **traceroute** commands, verify that data traffic between the local VPRN on PE-1 and the remote VPRN on PE-2 uses the modified Flex-Algorithm.

For IPv4 data traffic:

```

*A:PE-1# ping router 6 172.16.222.1
PING 172.16.222.1 56 data bytes
64 bytes from 172.16.222.1: icmp_seq=1 ttl=64 time=2.77ms.
---snip---
---- 172.16.222.1 PING Statistics ----
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min = 1.86ms, avg = 2.44ms, max = 2.77ms, stddev = 0.329ms

*A:PE-1# traceroute router 6 172.16.222.1
traceroute to 172.16.222.1, 30 hops max, 40 byte packets
 1 172.16.222.1 (172.16.222.1)  3.09 ms  3.08 ms  2.33 ms

```

For IPv6 data traffic:

```

*A:PE-1# ping router 6 2001:db8:222::1
PING 2001:db8:222::1 56 data bytes
64 bytes from 2001:db8:222::1 icmp_seq=1 hlim=64 time=2.53ms.
---snip---
---- 2001:db8:222::1 PING Statistics ----
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min = 1.88ms, avg = 2.31ms, max = 2.86ms, stddev = 0.356ms

*A:PE-1# traceroute router 6 2001:db8:222::1
traceroute to 2001:db8:222::1, 30 hops max, 60 byte packets
 1 2001:db8:222::1 (2001:db8:222::1)  2.74 ms  2.40 ms  2.64 ms

```

For the example with the TE-metric-based flexible algorithm, the same set of **show** commands as for the example with the delay-based flexible algorithm shows that:

- the metric type changes to *te-metric* (from *delay*)
- the TE APP LINK ATTR sub-TLVs contain a **TE Metric** value

- the applicable metric for the route between the local VPRN on PE-1 and the remote VPRN on PE-2 changes to 1110 (from 111 microseconds), corresponding with the sum of the metric values along the path PE-1, P-3, P-4, PE-2, as shown in the following **show router isis 0 topology flex-algo 128 detail** command output examples.

On PE-1:

```
*A:PE-1# show router isis 0 topology flex-algo 128 detail
```

```
=====
Rtr Base ISIS Instance 0 Flex-Algo 128 Topology Table
=====
```

```
-----
IS-IS IP paths (MT-ID 0), Level 2
-----
```

```
Node      : PE-2.00
Nexthop   : P-3
Interface : int-PE-1-P-3
SNPA      : none
Metric    : 1110

Node      : P-3.00
Nexthop   : P-3
Interface : int-PE-1-P-3
SNPA      : none
Metric    : 10

Node      : P-4.00
Nexthop   : P-3
Interface : int-PE-1-P-3
SNPA      : none
Metric    : 110
=====
```

On P-3:

```
*A:P-3# show router isis 0 topology flex-algo 128 detail
```

```
=====
Rtr Base ISIS Instance 0 Flex-Algo 128 Topology Table
=====
```

```
-----
IS-IS IP paths (MT-ID 0), Level 2
-----
```

```
---snip---
Node      : PE-2.00
Nexthop   : P-4
Interface : int-P-3-P-4
SNPA      : none
Metric    : 1100

Node      : P-4.00
Nexthop   : P-4
Interface : int-P-3-P-4
SNPA      : none
Metric    : 100
=====
```

On P-4:

```
*A:P-4# show router isis 0 topology flex-algo 128 detail
```

```
=====
Rtr Base ISIS Instance 0 Flex-Algo 128 Topology Table
=====
```

```
-----  
IS-IS IP paths (MT-ID 0),   Level 2  
-----  
---snip---  
Node       : PE-2.00  
Nexthop    : PE-2  
Interface  : int-P-4-PE-2  
SNPA      : none                Metric      : 1000  
---snip---  
=====
```

## Conclusion

The Flex-Algorithm for SRv6-based VPRNs feature allows the computation of constraint-based paths across an SRv6-enabled network, based on metrics other than the default IGP metrics. This allows carrying data traffic over an end-to-end path that is optimized using the best suited metric (IGP, delay, or TE).

## Inter-AS VPRN Model B

This chapter describes the Inter-AS VPRN Model B.

Topics in this chapter include:

- [Applicability](#)
- [Overview](#)
- [Configuration](#)
- [Conclusion](#)

### Applicability

This chapter was initially written for SR OS Release 15.0.R8, but the CLI in the current edition is based on SR OS Release 22.2.R1.

### Overview

An inter-AS Virtual Private Routed Network (VPRN) contains sites that are connected to different Autonomous Systems (ASs). Inter-AS is typically used either to provide extended reach through a partnership/trust agreement, as an interim means to interconnect ASs following acquisition, or because of the internal organization of a single Service Provider (SP). Three models for interconnecting ASs are defined in RFC 4364, labeled model A, B, and C. This chapter describes model B.

Inter-AS VPRN model B encompasses EBGP redistributing VPN-IPv4 and VPN-IPv6 routes between neighboring ASs. An Autonomous System Border Router (ASBR) learns VPN routes from within its AS using IBGP, potentially as a client of a Route Reflector (RR), then uses EBGP to redistribute those labeled VPN routes to its adjacent ASBR.

When redistributing the routes into EBGP, the ASBR imposes next-hop-self on the VPN-IPv4 and VPN-IPv6 update messages and generates its own label value when it advertises the update message upstream. Therefore, the ASBR programs a label-swap entry in its FIB and forwards traffic to the neighboring ASBR using a single-level label stack (the VPN label).

A key property of model B is that it eliminates the need for per-VPRN configuration on the ASBRs. However, both ASBRs must have a mechanism to implicitly learn all VPN prefixes within their local AS and selectively advertise some of those prefixes to the neighboring ASBR.

**Figure 11: Inter-AS VPRN Model B control and data plane example** shows an example of the control plane and corresponding data plane used in model B, where MPLS is used for transport in both ASs. CE-1 is attached to PE-1 in AS 64496 and advertises prefix 172.31.100.0/24, which is propagated between neighboring ASBRs to PE-2 in AS 64510 and upstream to CE-2.

The IP traffic originating from CE-2 and received by PE-2 is received on the VRF interface of VPRN 100 and encapsulated using a two-level label stack; the inner label is the VPN label (300) and the outer label is the LDP transport label used for reaching the local ASBR-2.

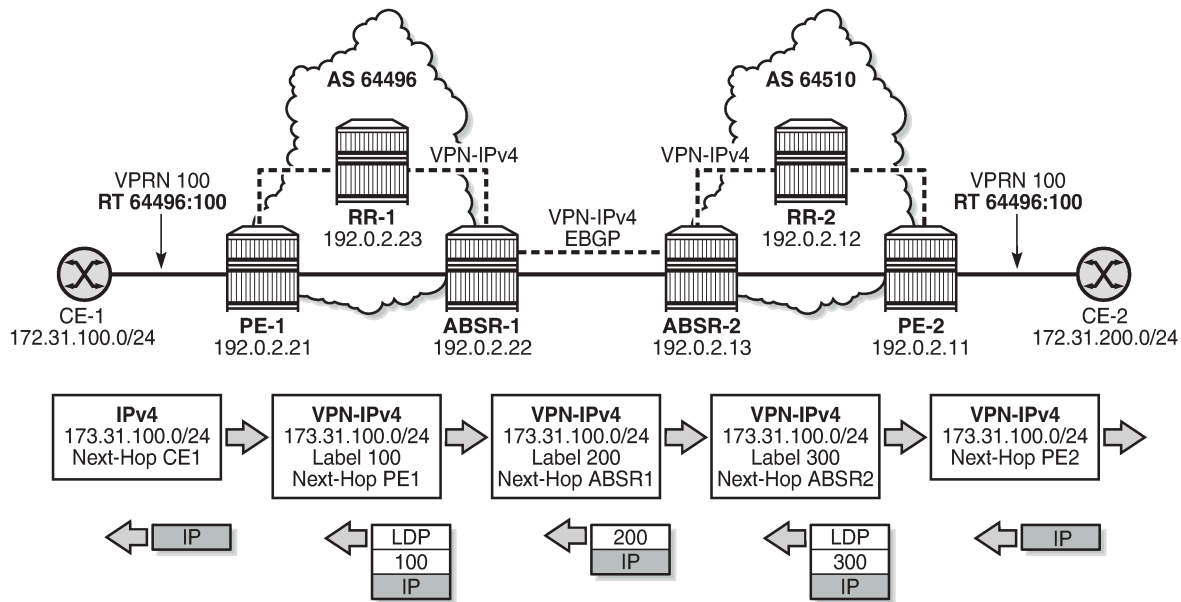


ASBR-2 passes the traffic to ASBR-1, removing the LDP transport label and swapping the VPN label (300) with its VPN label (200), resulting in a single-level label stack.

In turn, ASBR-1 swaps the received VPN label (200) with another VPN label (100) and adds an LDP transport label to reach PE-1.

Finally, PE-1 removes the VPN label and delivers the unlabeled IP traffic to CE-1.

Figure 11: Inter-AS VPRN Model B control and data plane example

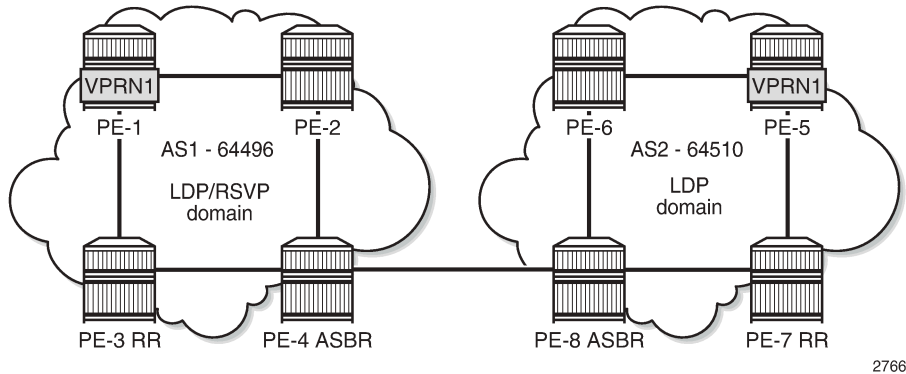


27660

## Configuration

In the example shown in [Figure 12: Inter-AS VPRN Model B topology](#), IS-IS is configured in each AS, and MP-IBGP sessions are established between the PEs and the RRs in AS 64496 and 64510, PE-3 and PE-7, respectively. LDP and RSVP-TE is used for transport in AS 64496, whereas AS 64510 uses LDP for its transport. An MP-EBGP session is established between ASBR PE-4 and ASBR PE-8.

Figure 12: Inter-AS VPRN Model B topology



27661

The initial configuration includes:

- Cards, MDAs, and ports
- Router interfaces
- IS-IS as IGP on all interfaces (alternatively, OSPF can be used), with traffic engineering enabled
- LDP and RSVP-TE configured in AS 64496, LDP configured in AS 64510
- IBGP configured in AS 64496, with PE-3 as RR for clients PE-1, PE-2, and PE-4
- IBGP configured in AS 64510, with PE-7 as RR for clients PE-5, PE-6, and PE-8

## Model B configuration

There are no specific requirements on PE routers or RRs for enabling inter-AS VPRN model B; only specific configurations are required on the ASBRs.

First, an ASBR must learn the VPN-IPv4 and VPN-IPv6 routes from the local AS and export these routes to the neighbor AS over an MP-EBGP session. This is achieved on each ASBR by declaring an IBGP group for peering with the local RR, and declaring an EBGP group for peering with the neighboring AS. The IBGP and EBGP groups have included the address family **vpn-ipv4**, **vpn-ipv6**, or both.

Additionally, import and export policies can be used to control the VPN-IPv4 and VPN-IPv6 routes exchanged. The latter requires the **vpn-apply-import** and **vpn-apply-export** commands for SR OS to match the prefixes of the VPN-IPv4 and VPN-IPv6 address families.

The use of the **next-hop-resolution** command is explained in the [Service configuration](#) section. The BGP configuration on ASBR PE-4 is as follows:

```
# on ASBR PE-4:
configure
  router Base
    bgp
      loop-detect discard-route
      enable-inter-as-vpn
      split-horizon
      next-hop-resolution
      labeled-routes
      transport-tunnel
```

```

        family vpn
            resolution-filter
                ldp # by default enabled for VPN routes
                rsvp
            exit
            resolution filter
        exit
    exit
exit
exit
group "vpn-eBGP"
    neighbor 192.168.48.2
        family vpn-ipv4 vpn-ipv6
        peer-as 64510
    exit
exit
group "vpn-iBGP"
    peer-as 64496
    neighbor 192.0.2.3
        family vpn-ipv4 vpn-ipv6
    exit
exit
no shutdown
exit
exit
exit

```

The configuration on ASBR PE-8 is similar.

Second, the **enable-inter-as-vpn** command enables the inter-AS functionality and causes the ASBR to store the received VPN-IPv4 routes in its RIB-In, even though it has no VRF that imports these routes. For a route to be considered valid, the ASBR still needs to resolve the next-hop of this route to a tunnel. The **enable-inter-as-vpn** command will also change the BGP next-hop of advertised and received VPN-IPv4/VPN-IPv6 routes. When a route is advertised to an EBGP peer, the BGP next-hop is changed to the local-address used for communicating with the EBGP peer. When a route is received from an EBGP peer and advertised to an IBGP peer, the BGP next-hop is changed to the local-address used for communicating with the IBGP peer.

The configuration of the MP-EBGP session between the ASBRs in the EBGP group allows the ASBR to forward labeled packets over its connection with its peer ASBR.

## MPLS LSP configuration

Two LSPs are needed between the end-to-end PEs (PE-1 and PE-5) to exchange service traffic bidirectionally, because LSPs are unidirectional. In AS 64496, this is achieved by configuring a first LSP from the service PE (PE-1) to the local ASBR (PE-4), and a second LSP back from the local ASBR (PE-4) toward the service PE (PE-1). In AS 64510, LDP is enabled on all interfaces; no RSVP LSPs are used.

In AS 64496, LDP and RSVP are enabled. The LSP (and path) from PE-1 to PE-4 runs via PE-3, as follows:

```

# on PE-1
configure
    router Base
        mpls
            path "path-PE-1-PE-3-PE-4"
                hop 10 192.168.13.2 strict
                hop 20 192.168.34.2 strict
            no shutdown

```

```

        exit
        lsp "lsp-PE-1-PE-4"
            to 192.0.2.4
            primary "path-PE-1-PE-3-PE-4"
            exit
            no shutdown
        exit
        no shutdown
    exit
exit
exit
exit

```

The LSP (and path) from PE-4 to PE-3 also runs via PE-3, as follows:

```

# on ASBR PE-4:
configure
router Base
    mpls
        path "path-PE-4-PE-3-PE-1"
            hop 10 192.168.34.1 strict
            hop 20 192.168.13.1 strict
            no shutdown
        exit
        lsp "lsp-PE-4-PE-1"
            to 192.0.2.1
            primary "path-PE-4-PE-3-PE-1"
            exit
            no shutdown
        exit
        no shutdown
    exit
exit
exit
exit

```

## Service configuration

VPRN 1 is configured on PE-1 and PE-5. Although the VPRN service IDs used in both ASs do not need to match, in an inter-AS VPRN model B context, the route targets (RTs) used in both ASs must be coordinated. The RT exported by the PE-1 VPRN 1 must be imported by the PE-5 VPRN 1, and vice versa. In this example, no specific **vrf-import** and **vrf-export** communities are used; the simplified method using a single **vrf-target** community is used instead.

To carry the customer data across AS 64496, tunnels must bind to a VPRN service with the **auto-bind-tunnel** command. Resolution is set to filter, indicating that SR OS must select a tunnel using the information defined in the **resolution-filter** context. The keywords **ldp** and **rsvp** in the resolution-filter context indicate that LDP or RSVP tunnels can be used, but SR OS prefers the RSVP tunnels because the preference for RSVP (7) is lower than the preference for LDP (9).

In AS 64496, the VPRN service on PE-1 is defined as follows:

```

# on PE-1:
configure
service
    vprn 1 name "VPRN1" customer 1 create
    interface "int-S1-1" create
        address 10.1.10.1/24
        ipv6
            address 2001:db8:1::1/120
    exit

```

```

        sap 1/2/1:1 create
        exit
    exit
    interface "int-S1-2" create
        address 10.1.11.1/24
        loopback
    exit
    bgp-ipvpn
        mpls
            auto-bind-tunnel
            resolution-filter
            ldp
            rsvp
            exit
            resolution filter
        exit
        route-distinguisher 64496:1
        vrf-target target:64496:1
        no shutdown
    exit
    exit
    no shutdown
exit

```

In AS 64510, the transport technology is LDP only, so the VPRN service in PE-5 auto-binds using LDP LSPs in the tunnel table to resolve VPN-IPv4 and VPN-IPv6 routes for which the vrf-target matches the vrf-target community value configured in PE-1, as follows:

```

# on PE-5 in AS 64510:
configure
    service
        vprn 1 name "VPRN1" customer 1 create
            description "VPN-1, counterpart is on PE-1"
            interface "int-S1-1" create
                address 10.1.50.1/24
                ipv6
                    address 2001:db8:1::5:1/120
            exit
            sap 1/2/1:1 create
            exit
        exit
        interface "int-S1-2" create
            address 10.1.51.1/24
            loopback
        exit
        bgp-ipvpn
            mpls
                auto-bind-tunnel
                resolution-filter
                ldp
                exit
                resolution filter
            exit
            route-distinguisher 64510:1
            vrf-target target:64496:1
            no shutdown
        exit
    exit
    no shutdown
exit

```

A second service is defined on PE-1 and PE-2 (VPRN 33), using loopback addresses 10.33.1.1/32 and 10.33.2.1/32 in PE-1 and PE-2, respectively. These addresses might appear in traces and commands later, but are of no concern because these are used for transporting intra-AS traffic.

For service traffic to flow in the PE-5 to PE-1 direction, ASBR PE-4 in AS 64496 must offer the possibility to use RSVP-TE tunnels when resolving a BGP next-hop for VPN services. Therefore, ASBR PE-4 must be explicitly configured, as follows:

```
# on ASBR PE-4:
configure
  router Base
    bgp
      next-hop-resolution
      labeled-routes
      transport-tunnel
      family vpn
      resolution-filter
      rsvp
    exit
      resolution filter
    exit
  exit
exit
exit
exit
exit
```

On ASBR PE-8 in AS 64510, no explicit configuration is required because resolving a BGP next-hop for VPN service to LDP tunnels is the default behavior.

## Verification

With the configurations from previous sections applied, PE-1 receives three VPN-IPv4 routes and one VPN-IPv6 route, as follows:

```
*A:PE-1# 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
-----
192.0.2.3
Def. Inst          64496    421   0 03h27m16s 3/3/3 (VpnIPv4)
                   424     0    1/1/1 (VpnIPv6)
-----
```

PE-1 received the following three VPN-IPv4 routes:

```
*A:PE-1# show router bgp routes vpn-ipv4

=====
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
Origin codes  : i - IGP, e - EGP, ? - incomplete
```

=====  
BGP VPN-IPv4 Routes  
=====

Flag	Network Nextthop (Router) As-Path	LocalPref Path-Id	MED IGP Cost Label
u*>i	64496:33:10.33.2.0/24 192.0.2.2 No As-Path	100 None	None 10 524283
u*>i	64510:1:10.1.50.0/24 192.0.2.4 64510	100 None	None 20 524279
u*>i	64510:1:10.1.51.0/24 192.0.2.4 64510	100 None	None 20 524279

-----  
Routes : 3  
=====

PE-1 received the following VPN-IPv6 route:

```
*A:PE-1# show router bgp routes vpn-ipv6
```

```
=====  
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
Origin codes  : i - IGP, e - EGP, ? - incomplete
```

=====  
BGP VPN-IPv6 Routes  
=====

Flag	Network Nextthop (Router) As-Path	LocalPref Path-Id	MED IGP Cost Label
u*>i	64510:1:2001:db8:1::5:0/120 ::ffff:192.0.2.4 64510	100 None	None 20 524278

-----  
Routes : 1  
=====

PE-1 has three LDP tunnels and one RSVP tunnel, and its tunnel table looks as follows:

```
*A:PE-1# show router tunnel-table
```

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

Destination Color	Owner	Encap	TunnelId	Pref	Nextthop	Metric
192.0.2.2/32	ldp	MPLS	65537	9	192.168.12.2	10
192.0.2.3/32	ldp	MPLS	65538	9	192.168.13.2	10
192.0.2.4/32	rsvp	MPLS	1	7	192.168.13.2	16777215
192.0.2.4/32	ldp	MPLS	65539	9	192.168.12.2	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 IPv4 routing table for VPRN 1 is as follows:

```
*A:PE-1# show router 1 route-table
=====
Route Table (Service: 1)
=====
Dest Prefix[Flags]
Next Hop[Interface Name]          Type   Proto   Age           Pref
                                   Metric
-----
10.1.10.0/24                       Local  Local   00h02m15s    0
int-S1-1                           0
10.1.11.0/24                       Local  Local   00h02m15s    0
int-S1-2                           0
10.1.50.0/24                       Remote BGP VPN 00h01m01s    170
192.0.2.4 (tunneled:RSVP:1)        16777215
10.1.51.0/24                       Remote BGP VPN 00h01m01s    170
192.0.2.4 (tunneled:RSVP:1)        16777215
-----
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
=====
```

The IPv4 addresses for VPRN 1 on PE-8 are 10.1.50.0/24 and 10.1.51.0/24, and are reachable through RSVP-TE tunnel 1 (*tunneled:RSVP:1*). The VPN label value for these prefixes is assigned and advertised by ASBR PE-4 and gets to PE-1 via the RR PE-3 in an MP-BGP update message. The 10.33.2.0/24 prefix belongs to a different service and is not relevant for model B because it is used for intra-AS traffic. The VPN-IPv4 routes received on PE-1 are as follows:

```
*A:PE-1# show router bgp neighbor 192.0.2.3 received-routes vpn-ipv4
=====
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
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  64496:33:10.33.2.0/24    100        None
      192.0.2.2            None        10
      No As-Path          524283
u*>i  64510:1:10.1.50.0/24    100        None
      192.0.2.4            None        20
      64510                524279
u*>i  64510:1:10.1.51.0/24    100        None
```



```

192.0.2.4          None          20
64510             524279
-----
Routes : 3
=====

```

The BGP next-hops for the VPN-IPv4 BGP address family are as follows. Service traffic for VPRN 33 uses the LDP tunnel to PE-2 carrying the intra-AS traffic, and service traffic for VPRN 1 uses the RSVP tunnel to ASBR PE-4 carrying the inter-AS traffic.

```

*A:PE-1# show router bgp next-hop vpn-ipv4
=====
BGP Router ID:192.0.2.1      AS:64496      Local AS:64496
=====

BGP VPN Next Hop
=====
VPN Next Hop              Owner
Autobind                  FibProg Reason
Labels (User-labels)     FlexAlgo Metric
Admin-tag-policy (strict-tunnel-tagging)
-----
192.0.2.2                 LDP
  ldp bgp                  Y
  -- (2)                   -- 10
  -- (-)
192.0.2.4                 RSVP
  ldp rsvp bgp             Y
  -- (2)                   -- 16777215
  -- (-)
-----
Next Hops : 2
=====

```

The IPv6 routing table for VPRN 1 is as follows:

```

*A:PE-1# show router 1 route-table ipv6
=====
IPv6 Route Table (Service: 1)
=====
Dest Prefix[Flags]      Type  Proto  Age      Pref
Next Hop[Interface Name]  Metric
-----
2001:db8:1::1:0/120    Local  Local  00h03m54s  0
  int-S1-1              0
2001:db8:1::5:0/120    Remote BGP VPN 00h02m40s 170
  192.0.2.4 (tunneled:RSVP:1) 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
=====

```

The VPN-IPv6 routes received on PE-1 are as follows:

```

*A:PE-1# show router bgp neighbor 192.0.2.3 received-routes vpn-ipv6
=====

```

```

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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP VPN-IPv6 Routes
=====
Flag  Network                               LocalPref  MED
      Nexthop (Router)                       Path-Id    IGP Cost
      As-Path                                Label
-----
u*>i  64510:1:2001:db8:1::5:0/120             100        None
      ::ffff:192.0.2.4                         None        20
      64510                                    524278
-----
Routes : 1
=====

```

The BGP next-hop for the VPN-IPv6 address family is as follows:

```

*A:PE-1# show router bgp next-hop vpn-ipv6
=====
BGP Router ID:192.0.2.1      AS:64496      Local AS:64496
=====
BGP VPN Next Hop
=====
VPN Next Hop                               Owner
Autobind                                   FibProg Reason
Labels (User-labels)                       FlexAlgo Metric
Admin-tag-policy (strict-tunnel-tagging)
-----
::ffff:192.0.2.4                           RSVP
  ldp rsvp bgp                               Y
  -- (2)                                     -- 16777215
  -- (-)
-----
Next Hops : 1
=====

```

The forwarding plane is programmed accordingly, as follows:

```

*A:PE-1# show router 1 fib 1 ipv4
=====
FIB Display
=====
Prefix [Flags]                               Protocol
NextHop
-----
10.1.10.0/24                                  LOCAL
  10.1.10.0 (int-S1-1)
10.1.11.0/24                                  LOCAL
  10.1.11.0 (int-S1-2)
10.1.50.0/24                                  BGP_VPN
  192.0.2.4 (VPRN Label:524279 Transport:RSVP LSP:1)
10.1.51.0/24                                  BGP_VPN
  192.0.2.4 (VPRN Label:524279 Transport:RSVP LSP:1)
-----

```

```

Total Entries : 4
-----
=====

*A:PE-1# show router 1 fib 1 ipv6

=====
FIB Display
=====
Prefix [Flags]                                Protocol
NextHop
-----
2001:db8:1::1:0/120                            LOCAL
  2001:db8:1::1:0 (int-S1-1)
2001:db8:1::5:0/120                            BGP_VPN
  192.0.2.4 (VPRN Label:524278 Transport:RSVP LSP:1)
-----
Total Entries : 2
-----
=====

```

SR OS uses a label-per-VRF mode of label distribution, meaning that the same label is used for different VPN-IPv4 and different VPN-IPv6 prefixes from the same VRF, which saves on MPLS label resources. In this example, the VPRN service label is 524279 for the VPN-IPv4 prefixes 10.1.50.0/24 and 10.1.51.0/24, and 524278 for VPN-IPv6 prefix 2001:db8:1::5:0/120.

The forwarding plane is also programmed with the outer label to be used for transport purposes. Two labels are present: 524282 assigned through RSVP, and 524284 assigned through LDP. Because RSVP takes precedence over LDP, the RSVP label is actively used, as follows:

```

*A: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/SID
NextHop
Lbl/SID (backup)                            Intf/Tunnel
NextHop (backup)
-----
192.0.2.4/32                                LDP          -
  524284
  192.168.12.2                               1/1/1:1000
192.0.2.4/32                                RSVP        1
  524282
  192.168.13.2                               1/1/2:1000
-----
Total Entries : 2
-----
=====

```

Traffic over VPRN 1 is generated using a ping command on PE-1 to the remote loopback address, as follows:

```
*A:PE-1# ping router 1 10.1.50.1
```

```

PING 10.1.50.1 56 data bytes
64 bytes from 10.1.50.1: icmp_seq=1 ttl=64 time=6.11ms.
64 bytes from 10.1.50.1: icmp_seq=2 ttl=64 time=6.13ms.
64 bytes from 10.1.50.1: icmp_seq=3 ttl=64 time=6.61ms.
64 bytes from 10.1.50.1: icmp_seq=4 ttl=64 time=6.00ms.
64 bytes from 10.1.50.1: icmp_seq=5 ttl=64 time=6.05ms.

---- 10.1.50.1 PING Statistics ----
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min = 6.00ms, avg = 6.18ms, max = 6.61ms, stddev = 0.220ms
    
```

On PE-1, the IPv4 VPRN 1 service traffic is pushed with VPN label 524279, followed by RSVP-TE transport label 524282. ASBR PE-4 removes the RSVP-TE transport label and swaps the internal (advertised) VPN label 524279 with the external VPN label 524280 received from ASBR PE-8. For IPv6 VPRN 1 traffic, VPN label 524278 is swapped by VPN label 524279. The inter-AS BGP labels stored by ASBR PE-4 are as follows:

```

*A:PE-4# show router bgp inter-as-label

=====
BGP Inter-AS labels
Flags: B - entry has backup, P - entry is promoted
=====
NextHop                Received      Advertised    Label
                        Label         Label         Origin
-----
192.0.2.1              524281       524282        Internal
192.0.2.1              524282       524281        Internal
192.0.2.1              524282       524280        Internal
192.0.2.2              524283       524277        Internal
192.168.48.2          524279       524278        External
192.168.48.2          524280       524279        External
-----
Total Labels allocated: 6
=====
    
```

The forward data flow (from AS 64496 to AS 64510) for VPRN 1 uses the labels for which the label origin is external. The VPN labels used for the backward data flow (from AS 64510 to 64496) uses the labels for which the label origin is internal.

For brevity, the commands to display and check VPN prefixes and labels used in AS 64510 are omitted.

By disabling (**shutdown**) both RSVP LSPs between PE-1 and ASBR PE-4 in AS 64496, both PE-1 and PE-4 will select LDP tunnels for resolving VPN BGP next-hops. Then, the route table for VPRN 1 is as follows, where *tunneled* indicates an LDP tunnel is used to reach the next hop:

```

*A:PE-1# show router 1 route-table

=====
Route Table (Service: 1)
=====
Dest Prefix[Flags]      Type   Proto   Age           Pref
  Next Hop[Interface Name]      Metric
-----
10.1.10.0/24            Local  Local   04h29m37s    0
  int-S1-1                0
10.1.11.0/24            Local  Local   04h29m37s    0
  int-S1-2                0
10.1.50.0/24           Remote BGP VPN 00h00m08s    170
  192.0.2.4 (tunneled)  20
10.1.51.0/24           Remote BGP VPN 00h00m08s    170
    
```

```

192.0.2.4 (tunneled) 20
-----
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
=====

```

Only LDP tunnels are available in PE-1 and ASBR PE-4, as follows:

```

*A:PE-1# show router tunnel-table
=====
IPv4 Tunnel Table (Router: Base)
=====
Destination      Owner      Encap TunnelId  Pref  Nexthop      Metric
  Color
-----
192.0.2.2/32     ldp        MPLS  65537    9    192.168.12.2  10
192.0.2.3/32     ldp        MPLS  65538    9    192.168.13.2  10
192.0.2.4/32     ldp        MPLS  65539    9    192.168.12.2  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
=====

```

```

*A:PE-4# show router tunnel-table
=====
IPv4 Tunnel Table (Router: Base)
=====
Destination      Owner      Encap TunnelId  Pref  Nexthop      Metric
  Color
-----
192.0.2.1/32     ldp        MPLS  65538    9    192.168.24.1  20
192.0.2.2/32     ldp        MPLS  65539    9    192.168.24.1  10
192.0.2.3/32     ldp        MPLS  65537    9    192.168.34.1  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
=====

```

The BGP next-hop for VPN-IPv4 traffic in PE-1 also indicates that, to reach PE-5 via PE-4, an LDP tunnel is used, as follows:

```

*A:PE-1# show router bgp next-hop vpn-ipv4
=====
BGP Router ID:192.0.2.1      AS:64496      Local AS:64496
=====

BGP VPN Next Hop
=====
VPN Next Hop      Owner
Autobind          FibProg Reason
Labels (User-labels)  FlexAlgo Metric

```

```

Admin-tag-policy (strict-tunnel-tagging)
-----
192.0.2.2
  ldp bgp                                Y      LDP
  -- (2)                                --    10
  -- (-)
192.0.2.4
  ldp rsvp bgp                           Y      LDP
  -- (2)                                --    20
  -- (-)
-----
Next Hops : 2
=====

```

The forwarding plane is programmed accordingly, as follows:

```

*A:PE-1# show router 1 fib 1 ipv4
=====
FIB Display
=====
Prefix [Flags]                                Protocol
NextHop
-----
10.1.10.0/24                                  LOCAL
  10.1.10.0 (int-S1-1)
10.1.11.0/24                                  LOCAL
  10.1.11.0 (int-S1-2)
10.1.50.0/24                                BGP_VPN
  192.0.2.4 (VPRN Label:524279 Transport:LDP)
10.1.51.0/24                                BGP_VPN
  192.0.2.4 (VPRN Label:524279 Transport:LDP)
-----
Total Entries : 4
=====

```

```

*A: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
Lbl/SID (backup)
NextHop (backup)
-----
192.0.2.2/32                                LDP          -
  524287
  192.168.12.2                               1/1/1:1000
192.0.2.3/32                                LDP          -
  524287
  192.168.13.2                               1/1/2:1000
192.0.2.4/32                                LDP        -
  524284
  192.168.12.2                               1/1/1:1000
-----
Total Entries : 3

```

The details for the LDP tunnel from PE-1 to PE-4 are as follows:

```
*A:PE-1# show router tunnel-table 192.0.2.4/32 detail
=====
Tunnel Table (Router: Base)
=====
Destination      : 192.0.2.4/32
NextHop         : 192.168.12.2
Tunnel Flags      : (Not Specified)
Age               : 00h12m46s
CBF Classes       : (Not Specified)
Owner           : ldp
Tunnel ID         : 65539
Tunnel Label    : 524284
Tunnel MTU        : 1556
Encap              : MPLS
Preference        : 9
Tunnel Metric     : 20
Max Label Stack   : 1
=====
Number of tunnel-table entries      : 1
Number of tunnel-table entries with LFA : 0
=====
```

On PE-1, the IPv4 traffic in VPRN 1 is pushed with VPN label 524279, followed by LDP transport label 524284. ASBR PE-4 removes the LDP transport label and swaps the internal (advertised) VPN label 524279 with the external VPN label 524280 received from ASBR PE-8. The inter-AS label mapping between the ASBRs remains unchanged.

On the directly connected interface between the ASBRs, nothing has changed; only a single MPLS label is used to carry the VPN data, as shown in the following capture:

With this configuration, all the VPN-IPv4 and VPN-IPv6 routes known to AS 64496 are advertised by ASBR PE-4 to AS 64510, even the VPN-IPv4 and VPN-IPv6 routes from other AS 64496 VPRN services that do not need to be distributed:

```
*A:PE-4# show router bgp neighbor 192.168.48.2 advertised-routes vpn-ipv4 brief
=====
BGP Router ID:192.0.2.4      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP VPN-IPv4 Routes
=====
Flag Network
-----
i    64496:1:10.1.10.0/24
i    64496:1:10.1.11.0/24
i    64496:33:10.33.1.0/24
i    64496:33:10.33.2.0/24
-----
Routes : 4
=====
```

As already indicated, the 10.33.1.0/24 and 10.33.2.0/24 prefixes belong to VPRN 33. This service exists on PE-1 and PE-2 only, and the corresponding customer traffic must be kept within AS 64496, so there is

no need to advertise these prefixes to the peer AS. The "exp-SVC-1" policy is defined at ASBR PE-4 to achieve this, as follows:

```
# on ASBR PE-4:
configure
router Base
  policy-options
  begin
  prefix-list "pfx-SVC-1"
    prefix 10.1.10.0/24 longer
    prefix 10.1.11.0/24 longer
    prefix 2001:db8:1::/96 longer
  exit
  policy-statement "exp-SVC-1"
  entry 10
    from
      prefix-list "pfx-SVC-1"
    exit
    action accept
  exit
  exit
  default-action drop
  exit
  exit
  commit
exit
exit
exit
```

The "exp-SVC-1" policy is applied to ASBR PE-4 as an export policy, but also import policies can be used to control which prefixes are exchanged. This additionally requires the **vpn-apply-export** (and the **vpn-apply-import**) command, and the change required at ASBR PE-4 is as follows:

```
# on ASBR PE-4:
configure
router Base
  autonomous-system 64496
  bgp
  group "vpn-eBGP"
  vpn-apply-export
  export "exp-SVC-1"
  neighbor 192.168.48.2
    family vpn-ipv4 vpn-ipv6
    peer-as 64510
  exit
  exit
  exit
exit
exit
exit
```

Therefore, the PE-4 ASBR will only advertise the VPN-IPv4 and VPN-IPv6 prefixes for VRPN 1, as follows:

```
*A:PE-4# show router bgp neighbor 192.168.48.2 advertised-routes vpn-ipv4 brief
=====
BGP Router ID:192.0.2.4      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
```



```
BGP VPN-IPv4 Routes
=====
Flag  Network
-----
i     64496:1:10.1.10.0/24
i     64496:1:10.1.11.0/24
-----
Routes : 2
=====
```

```
*A:PE-4# show router bgp neighbor 192.168.48.2 advertised-routes vpn-ipv6 brief
=====
BGP Router ID:192.0.2.4      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP VPN-IPv6 Routes
=====
Flag  Network
-----
i     64496:1:2001:db8:1::1:0/120
-----
Routes : 1
=====
```

## Conclusion

Inter-AS VPRN model B offers service providers a way to interconnect IPv4 and IPv6 VPN sites across different ASs, avoiding the need for dedicated services in the ASBR, which would otherwise consume valuable resources in the ASBR. Model B is useful for scenarios where model C does not apply; for example, when there is no trust agreement with the peer AS, so that exchanging PE system addresses with that peer is not permitted or does not make sense.

# Inter-AS VPRN Model B Using MPLS over UDP

This chapter describes Inter-AS VPRN Model B using MPLS over UDP.

Topics in this chapter include:

- [Applicability](#)
- [Overview](#)
- [Configuration](#)
- [Conclusion](#)

## Applicability

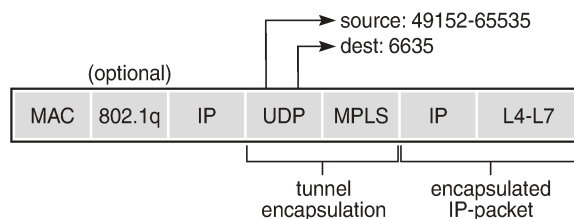
This chapter was initially written based on SR OS Release 15.0.R8, but the CLI in the current edition corresponds to SR OS Release 23.3.R3.

## Overview

MPLS over UDP encapsulates MPLS packets into a UDP tunnel that can be transported by any IP network, and is defined in RFC 7510.

With MPLS over UDP, an outer IPv4/UDP or IPv6/UDP header encapsulates the inner MPLS label stack and message body; see [Figure 13: IP over MPLS over UDP packet format](#). In the UDP header, the destination UDP port 6635 identifies the MPLS over UDP format to the egress PE and the source port number in the range from 49152 to 65535 is a source of entropy, because it is based on an ECMP hash calculation by the ingress PE. The entropy in the IP/UDP header ensures that ECMP uses all the available parallel paths between the tunnel endpoints.

Figure 13: IP over MPLS over UDP packet format



27658

The MPLS over UDP implementation in SR OS has the following characteristics:

- The UDP checksum is set to 0 on transmission and ignored on reception.
- MPLS over UDP packets are processed only if they arrive with an IP destination address matching the system IP address.

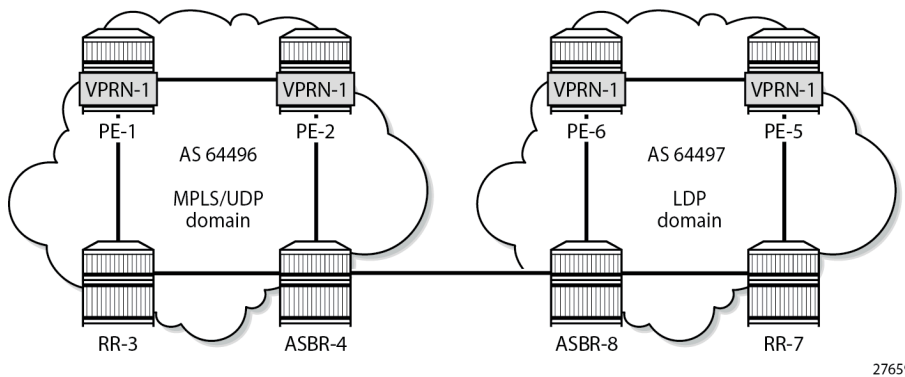
- MPLS over UDP packets are originated with the Don't Fragment (DF) bit set in the outer IP header.
- Reassembly of received MPLS over UDP packets is not supported.

BGP next-hops can be resolved to UDP-based tunnels for L3 VPN and EVPN services, which is useful, for example, in data center (DC) environments where IP is prevalent. This chapter describes the use and configuration of MPLS over UDP in an inter-AS VPRN model B environment. See the [Inter-AS VPRN Model B](#) chapter for more information.

## Configuration

**Figure 14: Inter-AS VPRN model B topology using MPLS over UDP in AS 64496** shows the topology used in this chapter, with two autonomous systems (ASs) and using inter-AS VPRN model B. VPRN 1 carries customer traffic between PE-1 and PE-2 in AS 64496, and PE-5 and PE-6 in AS 64497. IS-IS is configured in AS 64496 and 64497, and MP-IBGP sessions are established between the PEs of each AS. An MP-EBGP session is established between autonomous system border routers (ASBRs) ASBR-4 and ASBR-8. RR-3 and RR-7 are the route reflectors (RRs) in AS 64496 and 64497, respectively, for the VPN-IPv4 and VPN-IPv6 address families. AS 64496 uses MPLS over UDP for its transport, whereas AS 64497 uses LDP.

*Figure 14: Inter-AS VPRN model B topology using MPLS over UDP in AS 64496*



27659

An inter-AS VPRN model B requires configuration of an MP-EBGP session with support for the VPN-IPv4 and VPN-IPv6 address families and enabling the inter-AS functionality on the ASBRs by the **enable-inter-as-vpn** command. No configuration is required to support inter-AS VPRN model B on the PEs and the RRs. See the [Inter-AS VPRN Model B](#) chapter for more information.

The initial configuration includes:

- cards, MDAs, and ports
- router interfaces
- IS-IS as IGP on all interfaces (alternatively, OSPF can be used), with traffic engineering enabled.
- LDP configured in AS 64497, but not in AS 64496
- IBGP configured in both ASs
- EBGP configured between ASBR-4 and ASBR-8
- RR-3 and RR-7 configured as RR for VPN-IPv4 and VPN-IPv6 in AS 64496 and AS 64497

## MPLS over UDP configuration

MPLS over UDP tunnels are UDP-based LSP tunnels that are created dynamically through a BGP import policy, where the action is **create-udp-tunnel**. This import policy is configured and used on PE-1, PE-2, and ASBR-4. MPLS over UDP tunnels are created when BGP receives an update message where the incoming route has a next hop which matches the import policy. This import policy **create-UDP-tunnel** is defined as follows:

```
# on PE-1, PE-2, ASBR-4:
configure
  router Base
    policy-options
      begin
        prefix-list "system-pfxs"
          prefix 192.0.2.0/24 prefix-length-range 32-32
        exit
        policy-statement "create-UDP-tunnel"
          entry 10
            from
              next-hop prefix-list "system-pfxs"
              family vpn-ipv4 vpn-ipv6
            exit
            action accept
              create-udp-tunnel
            exit
          exit
        exit
      exit
    exit
  commit
```

This policy is applied to the IBGP sessions on which the VPN routes are exchanged; on PE-1 and PE-2, this is in the *IBGP-vpn* group, as follows:

```
# on PE-1, PE-2:
configure
  router Base
    autonomous-system 64496
    bgp
      loop-detect discard-route
      split-horizon
      next-hop-resolution
      labeled-routes
        transport-tunnel
          family vpn
            resolution-filter
              no ldp
              no bgp
              udp
            exit
          resolution filter
        exit
      exit
    exit
  exit
  group "IBGP-vpn"
    vpn-apply-import
    import "create-UDP-tunnel"
    peer-as 64496
    neighbor 192.0.2.3
      family vpn-ipv4 vpn-ipv6
    exit
```

```
exit
```

On RR-3, BGP is configured as follows:

```
# on RR-3:
configure
  router Base
    autonomous-system 64496
    bgp
      loop-detect discard-route
      split-horizon
      group "IBGP-vpn"
        family vpn-ipv4 vpn-ipv6
        cluster 192.0.2.3
        peer-as 64496
        neighbor 192.0.2.1
        exit
        neighbor 192.0.2.2
        exit
        neighbor 192.0.2.4
        exit
      exit
```

For ASBR-4 to advertise VPRN routes to the peer AS, ASBR-4 must know the VPRN routes used within the AS, so it peers with RR-3 through IBGP, using the *IBGP-vpn* group. Although the **enable-inter-as-vpn** command enables inter-AS VPN model B, the VPN-IPv4 and VPN-IPv6 address family must also resolve to MPLS over UDP tunnels, so the BGP configuration on ASBR-4 is as follows:

```
# on ASBR-4:
configure
  router Base
    autonomous-system 64496
    bgp
      loop-detect discard-route
      enable-inter-as-vpn # inter-AS VPRN model B
      split-horizon
      next-hop-resolution
        labeled-routes
          transport-tunnel
            family vpn
              resolution-filter
                no ldp
                no bgp
                udp
              exit
            resolution filter
          exit
        exit
      exit
    exit
  exit
  group "EBGP-vpn"
    neighbor 192.168.48.2
    family vpn-ipv4 vpn-ipv6
    peer-as 64497
  exit
  group "IBGP-vpn"
    vpn-apply-import
    import "create-UDP-tunnel"
    peer-as 64496
    neighbor 192.0.2.3
    family vpn-ipv4 vpn-ipv6
```

```
exit
exit
```

## Service configuration

VPRN "VPRN-1" is configured on PE-1, PE-2, PE-5, and PE-6. Although the VPRN service names and IDs used in both ASs do not have to match, when inter-AS VPRN model B applies, the route targets (RTs) used in both ASs must be coordinated. The RT exported by the PE-1 VPRN must be imported by PE-2, PE-5, and PE-6, and vice versa. In this example, no specific VRF import and VRF export policies are used; the **vrf-target** command is used instead.

To carry the customer data across AS 64496, the MPLS over UDP tunnels must bind to a VPRN service with the **auto-bind-tunnel** command. The resolution is set to **filter**, indicating that SR OS must select a tunnel using the information defined in the **resolution-filter** context. The **udp** keyword in the **resolution-filter** context indicates that MPLS over UDP tunnels can be used. If **resolution** is set to **any**, the resolution filter is ignored, and a tunnel from the tunnel table manager (TTM) is selected, based on availability and preference.

In AS 64496, the service "VPRN-1" on PE-1 is defined as follows. The configuration of service "VPRN-1" on PE-2 is similar.

```
# on PE-1:
configure
service
  vprn 1 name "VPRN-1" customer 1 create
  interface "int-S1-1" create
    address 10.1.10.1/24
    ipv6
      address 2001:db8:1::1/120
    exit
  sap 1/1/c4/1:1 create
  exit
  bgp-ipvpn
  mpls
    auto-bind-tunnel
    resolution-filter
      no ldp
      no bgp
      udp
    exit
    resolution filter
  exit
  route-distinguisher 64496:1
  vrf-target target:64496:1
  no shutdown
  exit
exit
no shutdown
```

The transport technology in AS 64497 is LDP, so service "VPRN-1" in PE-5 and PE-6 auto-binds to LDP LSPs in the tunnel table, to resolve VPN-IPv4 routes for which the VRF target matches the VRF target community value configured in PE-1 and PE-2, as follows. The configuration of "VPRN-1" on PE-6 is similar.

```
# on PE-5:
configure
```

```

service
  vprn 1 name "VPRN-1" customer 1 create
  interface "int-S1-1" create
    address 10.1.50.1/24
    ipv6
      address 2001:db8:1::5:1/120
    exit
    sap 1/1/c4/1:1 create
    exit
  exit
  bgp-ipvpn
    mpls
      auto-bind-tunnel
      resolution-filter
      ldp
        no bgp
      exit
      resolution filter
    exit
    route-distinguisher 64497:1
    vrf-target target:64496:1 # same value as on PE-1, PE-2
    no shutdown
  exit
exit
no shutdown

```

## Verification

With the configurations from previous subsections applied, so that PE-1, PE-2, PE-5, and PE-6 have a service instance of "VPRN-1" running, PE-1 receives three VPN-IPv4 and three VPN-IPv6 prefixes, as follows:

```

*A:PE-1# 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
-----
192.0.2.3
Def. Inst          64496    19   0 00h04m24s 3/3/1 (VpnIPv4)
                   17     0   3/3/1 (VpnIPv6)
-----

```

PE-1 creates two UDP tunnels, and its tunnel table is as follows. The UDP tunnel to 192.0.2.2 is used for intra-AS customer traffic; the UDP tunnel to 192.0.2.4 is used for inter-AS customer traffic.

```

*A:PE-1# show router tunnel-table
=====
IPv4 Tunnel Table (Router: Base)
=====
Destination          Owner      Encap TunnelId  Pref  Nexthop      Metric
-----

```

```

Color
-----
192.0.2.2/32      udp      MPLS  786434  254    192.168.12.2  10
192.0.2.4/32      udp      MPLS  786435  254    192.168.12.2  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
=====

```

These tunnels are used as BGP next-hops for the VPN-IPv4 (and VPN-IPv6) routes, as follows:

```

*A:PE-1# show router bgp next-hop vpn-ipv4
=====
BGP Router ID:192.0.2.1      AS:64496      Local AS:64496
=====

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      UDP
udp            Y
-- (2)        --      10
-- (N)        --      00h05m17s
192.0.2.4      UDP
udp            Y
-- (2)        --      20
-- (N)        --      00h04m45s
-----
Next Hops : 2
=====

```

The IPv4 and IPv6 routing tables for VPRN 1 are as follows:

```

*A:PE-1# show router 1 route-table ipv4
=====
Route Table (Service: 1)
=====
Dest Prefix[Flags]      Type  Proto  Age      Pref
Next Hop[Interface Name] Metric
-----
10.1.10.0/24           Local Local  00h05m54s  0
int-S1-1               0
10.1.20.0/24           Remote BGP VPN 00h05m31s  170
192.0.2.2 (tunneled:UDP) 10
10.1.50.0/24           Remote BGP VPN 00h04m59s  170
192.0.2.4 (tunneled:UDP) 20
10.1.60.0/24           Remote BGP VPN 00h04m59s  170
192.0.2.4 (tunneled:UDP) 20
-----
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

```



```

=====
*A:PE-1# show router 1 route-table ipv6
=====
IPv6 Route Table (Service: 1)
=====
Dest Prefix[Flags]                               Type  Proto  Age           Pref
  Next Hop[Interface Name]                       Metric
-----
2001:db8:1::1:0/120                             Local  Local  00h05m53s    0
  int-S1-1                                       0
2001:db8:1::2:0/120                             Remote BGP VPN 00h05m31s   170
  192.0.2.2 (tunneled:UDP)                       10
2001:db8:1::5:0/120                             Remote BGP VPN 00h04m59s   170
  192.0.2.4 (tunneled:UDP)                       20
2001:db8:1::6:0/120                             Remote BGP VPN 00h04m59s   170
  192.0.2.4 (tunneled:UDP)                       20
-----
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
=====

```

In this case, all VPRN remote prefixes are reachable through an MPLS over UDP tunnel (tunneled:UDP). The VPN label values for these prefixes are assigned and advertised by ASBR-4 and get to PE-1 via RR-3 in an MP-BGP update message, and can be displayed as follows:

```

=====
*A:PE-1# show router bgp neighbor 192.0.2.3 received-routes vpn-ipv4
=====
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
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  64496:1:10.1.20.0/24                    100        None
      192.0.2.2                             None        10
      No As-Path                             524287
u*>i  64497:1:10.1.50.0/24                    100        None
      192.0.2.4                             None        20
      64497                                  524286
u*>i  64497:1:10.1.60.0/24                    100        None
      192.0.2.4                             None        20
      64497                                  524287
-----
Routes : 3
=====

```

The prefixes 10.1.50.0/24 and 10.1.60.0/24 have ASBR-4 as next hop: prefix 10.1.50.0/24 gets VPRN label 524286 and prefix 10.1.60.0/24 gets VPRN label 524287.

```
*A:PE-1# show router bgp neighbor 192.0.2.3 received-routes vpn-ipv6
=====
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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP VPN-IPv6 Routes
=====
Flag  Network                               LocalPref  MED
      Nexthop (Router)                     Path-Id    IGP Cost
      As-Path                               Path-Id    Label
-----
u*>i  64496:1:2001:db8:1::2:0/120           100        None
      ::ffff:192.0.2.2                     None       10
      No As-Path                           None       524287
u*>i  64497:1:2001:db8:1::5:0/120           100        None
      ::ffff:192.0.2.4                     None       20
      64497                                 None       524285
u*>i  64497:1:2001:db8:1::6:0/120           100        None
      ::ffff:192.0.2.4                     None       20
      64497                                 None       524284
-----
Routes : 3
=====
```

The prefixes 2001:db8:1::5:0/120 and 2001:db8:1::6:0/120 have ASBR-4 as next hop: prefix 2001:db8:1::5:0/120 gets VPRN label 524285 and prefix 2001:db8:1::6:0/120 gets VPRN label 524284.

The forwarding plane is programmed accordingly, as follows:

```
*A:PE-1# show router 1 fib 1 ipv4
=====
FIB Display
=====
Prefix [Flags]                               Protocol
NextHop
-----
10.1.10.0/24                                  LOCAL
  10.1.10.0 (int-S1-1)
10.1.20.0/24                                  BGP_VPN
  192.0.2.2 (VPRN Label:524287 Transport:UDP)
10.1.50.0/24                                  BGP_VPN
  192.0.2.4 (VPRN Label:524286 Transport:UDP)
10.1.60.0/24                                  BGP_VPN
  192.0.2.4 (VPRN Label:524287 Transport:UDP)
-----
Total Entries : 4
=====
```

```
*A:PE-1# show router 1 fib 1 ipv6
=====
FIB Display
```

```

=====
Prefix [Flags]                                Protocol
NextHop
-----
2001:db8:1::1:0/120                          LOCAL
  2001:db8:1::1:0 (int-S1-1)
2001:db8:1::2:0/120                          BGP_VPN
  192.0.2.2 (VPRN Label:524287 Transport:UDP)
2001:db8:1::5:0/120                          BGP_VPN
  192.0.2.4 (VPRN Label:524285 Transport:UDP)
2001:db8:1::6:0/120                          BGP_VPN
  192.0.2.4 (VPRN Label:524284 Transport:UDP)
-----
Total Entries : 4
=====

```

Traffic over the VPRN is generated using a **ping** command on PE-1 to the remote address, as follows:

```

*A:PE-1# ping router 1 10.1.50.1 source 10.1.10.1
PING 10.1.50.1 56 data bytes
64 bytes from 10.1.50.1: icmp_seq=1 ttl=64 time=6.06ms.
64 bytes from 10.1.50.1: icmp_seq=2 ttl=64 time=6.16ms.
64 bytes from 10.1.50.1: icmp_seq=3 ttl=64 time=6.12ms.
64 bytes from 10.1.50.1: icmp_seq=4 ttl=64 time=5.89ms.
64 bytes from 10.1.50.1: icmp_seq=5 ttl=64 time=6.00ms.

---- 10.1.50.1 PING Statistics ----
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min = 5.89ms, avg = 6.05ms, max = 6.16ms, stddev = 0.094ms

```

In contrast with the traditional inter-AS VPRN model B, where customer traffic is pushed with a VPN label followed by a transport label, now customer traffic to destination 10.1.50.1 is pushed with VPN label 524286, followed by an IP/UDP header with IP SA 192.0.2.1 and IP DA 192.0.2.4 and with UDP source and destination port.

The interconnection between the ASBRs carries the VPN data with a single MPLS label, so ASBR-4 removes the IP/UDP header and swaps the VPN label 524286 with the VPN label 524282 received from ASBR-8. The inter-AS label mapping on ASBR-4 is as follows:

```

*A:ASBR-4# show router bgp inter-as-label

=====
BGP Inter-AS labels
Flags: B - entry has backup, P - entry is promoted
=====
NextHop                Received      Advertised    Label
                        Label         Label         Origin
-----
192.0.2.1              524287       524283       Internal
192.0.2.1              524287       524281       Internal
192.0.2.2              524287       524282       Internal
192.0.2.2              524287       524280       Internal
192.168.48.2          524280       524285       External
192.168.48.2          524281       524284       External
192.168.48.2          524282       524286       External
192.168.48.2          524283       524287       External
-----
Total Labels allocated: 8
=====

```

The forward data flow from PE-1 to PE-5 for VPRN 1 uses the labels for which the label origin is external. The VPN labels used for the backward data flow use the labels for which the label origin is internal.

For brevity, the commands to display and check VPN prefixes and labels in AS 64497 are omitted.

Customer traffic destined to the VPN routes received by ASBR-4 can be sent to the correct destination because ASBR-4 has the relevant BGP next-hops resolved to UDP tunnels, as follows:

```
*A:ASBR-4# show router bgp next-hop vpn-ipv4
=====
BGP Router ID:192.0.2.4      AS:64496      Local AS:64496
=====
BGP VPN Next Hop
=====
VPN Next Hop      Owner
Autobind          FibProg Reason
Labels (User-labels) FlexAlgo Metric
Admin-tag-policy (strict-tunnel-tagging) Last Mod.
-----
192.0.2.1
  udp              Y          UDP
  -- (2)          --         20
  -- (N)          --         00h07m47s
192.0.2.2
  udp              Y          UDP
  -- (2)          --         10
  -- (N)          --         00h07m47s
192.168.48.2
  udp              Y          LOCAL
  -- (2)          --         0
  -- (N)          --         00h07m51s
-----
Next Hops : 3
=====
```

The forwarding plane is programmed accordingly, as follows:

```
*A:ASBR-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)
-----
192.0.2.1/32     UDP          -
-
  192.168.24.1   1/1/c2/1:1000
192.0.2.2/32     UDP          -
-
  192.168.24.1   1/1/c2/1:1000
-----
Total Entries : 2
=====
```

Changing the BGP next-hop resolution for auto-binding the BGP next-hop on PE-1, and for VPN next-hop resolution on ASBR-4 from **resolution filter** to **resolution any**, does not lead to the tunnel being changed, but to a change of the allowed tunnel types for auto-bind. On PE-1 and PE-2, SR OS selects UDP as the only viable tunnel type, because no tunnels of type LDP, RSVP, SR-ISIS, SR-OSPF, GRE and so on are available:

```
*A:PE-1# show router bgp next-hop vpn-ipv4
=====
BGP Router ID:192.0.2.1      AS:64496      Local AS:64496
=====

BGP VPN Next Hop
=====
VPN Next Hop
Autobind                      FibProg  Owner
Labels (User-labels)         FlexAlgo Reason
Admin-tag-policy (strict-tunnel-tagging)  Last Mod.
-----
192.0.2.2
  ldp rsvp sr-isis sr-ospf gre bgp sr-te udp sr-p Y
  olicy rib-api mpls-fwd-policy sr-ospf3
  -- (2)                      --      10
  -- (N)                      --      00h00m10s
192.0.2.4
  ldp rsvp sr-isis sr-ospf gre bgp sr-te udp sr-p Y
  olicy rib-api mpls-fwd-policy sr-ospf3
  -- (2)                      --      20
  -- (N)                      --      00h00m10s
-----
Next Hops : 2
=====
```

```
*A:ASBR-4# show router bgp next-hop vpn-ipv4
=====
BGP Router ID:192.0.2.4      AS:64496      Local AS:64496
=====

BGP VPN Next Hop
=====
VPN Next Hop
Autobind                      FibProg  Owner
Labels (User-labels)         FlexAlgo Reason
Admin-tag-policy (strict-tunnel-tagging)  Last Mod.
-----
192.0.2.1
  ldp rsvp sr-isis sr-ospf bgp sr-te udp sr-polic Y
  y rib-api mpls-fwd-policy sr-ospf3
  -- (2)                      --      20
  -- (N)                      --      00h00m19s
192.0.2.2
  ldp rsvp sr-isis sr-ospf bgp sr-te udp sr-polic Y
  y rib-api mpls-fwd-policy sr-ospf3
  -- (2)                      --      10
  -- (N)                      --      00h00m19s
192.168.48.2
  ldp rsvp sr-isis sr-ospf bgp sr-te udp sr-polic Y
  y rib-api mpls-fwd-policy sr-ospf3
  -- (2)                      --      0
-----
```

```
-- (N)                                00h00m19s
-----
Next Hops : 3
=====
```

## Conclusion

VPRN services support the resolution of VPN-IPv4 and VPN-IPv6 BGP next-hops to MPLS over UDP tunnels. MPLS over UDP tunnels are useful in IP-based fabric networks, such as DCs. SR OS supports inter-AS model B for any type of MPLS-based tunnels, including MPLS over UDP.

# Inter-AS VPRN Model C

This chapter provides information about virtual private routed network (VPRN) inter-autonomous system (AS) virtual private network (VPN) model C.

Topics in this chapter include:

- [Applicability](#)
- [Overview](#)
- [Configuration](#)
- [Conclusion](#)

## Applicability

This chapter was initially written for SR OS Release 7.0. The CLI in the current edition corresponds to SR OS Release 22.2.R1. There are no prerequisites for this configuration.

## Overview

### Introduction

Section 10 of RFC 4364, *BGP/MPLS IP Virtual Private Networks (VPNs)*, describes three potential methods for service providers to interconnect their IP-VPN (Internet Protocol — Virtual Private Network) backbones in order to provide an end-to-end MPLS-VPN where one or more sites of the VPN are connected to different service provider autonomous systems (ASs). The purpose of this chapter is to describe the configuration and troubleshooting for inter-AS VPN model C.

In this architecture, VPN prefixes are neither held, nor re-advertised by the Autonomous System Border Router — Provider Edge (ASBR-PE) routers, which makes Model C more scalable than Model B (where the only prefixes exchanged between ASs are VPN-IPv4). In Model C, the only MPLS data plane resources consumed in the ASBRs are for infrastructure addresses of PEs and RRs rather than VPN prefixes.

In this example, an export policy is configured to ensure that the nodes advertise their system IP addresses (IPv4 /32 addresses) in labeled BGP to all their BGP peers within the AS. Therefore, the ASBR-PE maintains labeled IPv4 /32 BGP routes to other PE routers within its own AS. These BGP routes are inactive, because for each destination within the AS, an IGP route exists which is preferred to BGP routes. The ASBR redistributes these inactive /32 IPv4 prefixes in external Border Gateway Protocol (EBGP) to the ASBR-PE in other service providers ASs, because **advertise-inactive** is configured in EBGP. No export policy is required in EBGP.

At the same time, the ASBR programs a label switch for the received and advertised BGP labels. The receiving ASBR advertises the received IP system prefixes to its IBGP peers (in this case, a Route Reflector (RR)) within their AS, and eventually, all PEs in the AS learn the system IP prefixes of the peer

AS. However, there is no need to learn the system IP address of the ASBRs in peer ASs, because they do not exchange customer VPN prefixes.

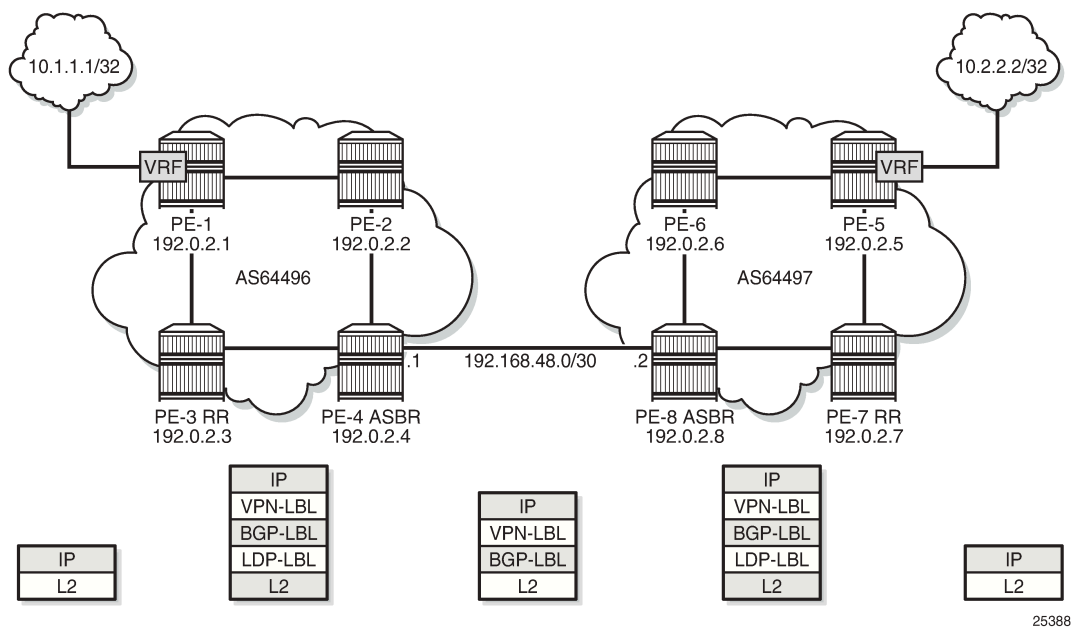
After the system IP addresses have been learned in the peer AS, it is possible for PE routers in different ASs to establish multi-hop Multi Protocol — external Border Gateway Protocol (MP-EBGP) sessions for address family VPN-IPv4 to each other in order to exchange customer VPN prefixes over those connections. The multihop sessions can be established between the RR in the ASs, but these RRs should not modify the next-hop attribute of the BGP update across the EBGP session.

A three-level label stack is imposed on the ingress PE. The bottom-level label is assigned by the egress PE (advertised in multi-hop MP-EBGP without next-hop override) and is commonly referred to as the VPN-label. The middle label is assigned by the local ASBR-PE and corresponds to the /32 route of the egress PE (in a different AS) using BGP-LBL (RFC 3107, *Carrying Label Information in BGP-4*). The top level label is the label assigned by the local ASBR-PE /32 loopback address, which is assigned by the IGP next-hop of the ingress PE. This label is referred to as the LDP-LBL.

**Figure 15: Inter-AS VPN Model C** illustrates this mechanism. The VPN-LBL is assigned by PE-5, the BGP-LBL is assigned by PE-4 and the LDP-LBL is also assigned by PE-4. The BGP-LBL is swapped in both ASBRs. The label stack contains three labels in each AS: VPN-LBL, BGP-LBL, and LDP-LBL) and two labels on the EBGP link between the ASs: VPN-LBL and BGP-LBL.

Note: This configuration that uses **advertise-inactive** is preferred to a configuration where the BGP routes are not exchanged within their AS and the ASBRs use an export policy with a prefix list for all local system prefixes to be advertised to the peer ASs. The routes for those prefixes are taken from the RTM, where these routes are not known via BGP, but via IS-IS. In that case, IS-IS routes are effectively redistributed into labeled BGP (which most operators do not want) and as a result, the ASBR is not programming a label switch for the BGP label. Furthermore, the label stack is asymmetrical: three labels in the originating AS (VPN-LBL, BGP-LBL, and LDP-LBL) and only two labels in the target AS (VPN-LBL, LDP-LBL), because the local routes are not known via labeled BGP in this scenario. This scenario is not explained in this chapter; only the preferred scenario with local labeled BGP routes in each AS is explained.

Figure 15: Inter-AS VPN Model C

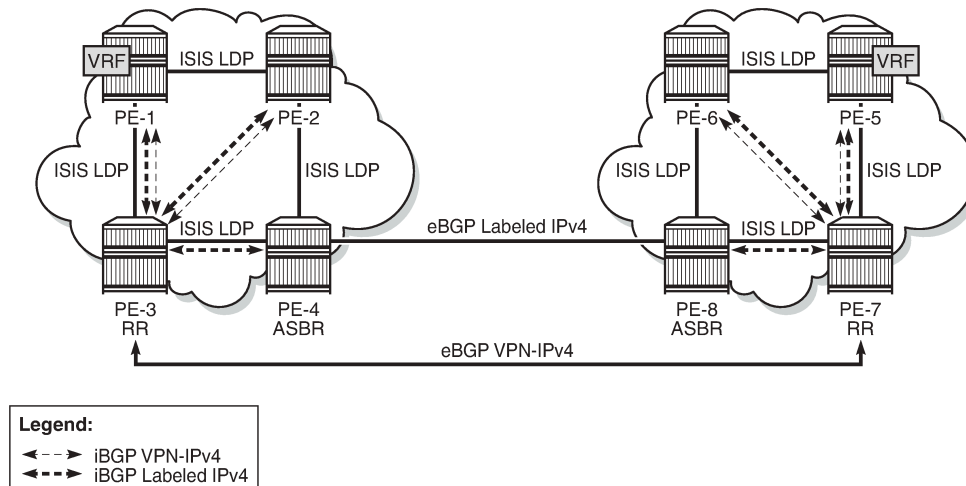




The VPN connectivity is established using Labeled VPN route exchange using MP-EBGP without next-hop override. The PE connectivity is established as follows.

EBGP PE /32 loopback leaking routing exchange using EBGP LBL (RFC 3107) at the ASBR-PE. The /32 PE routes learned from the other AS through the ASBR-PE are further distributed into the local AS using IBGP and optionally through Route Reflectors (RRs). This model uses a three label stack and is referred to as Model C. Resilience for ASBR-PE failures depends on BGP.

Figure 16: Protocol overview



25389

Figure 16: Protocol overview shows the protocols used when implementing Inter-AS Model C. Inside each AS, there is an IS-IS adjacency and a link LDP session between each pair of adjacent nodes. As an alternative, OSPF can be used as IGP. There is also an IBGP session between each PE and the RR. The address family is both VPN-IPv4 for the exchange of customer VPN prefixes and Labeled IPv4 for the exchange of labeled IPv4 prefixes. Between the RR and the ASBR, only Labeled IPv4 is required, because the ASBR does not exchange any customer VPN prefixes. When no RR is used, a full mesh of IBGP sessions can be established in each AS.

Between the ASBRs, there is an EBGP session for the exchange of labeled IPv4 prefixes. The ASBRs override the next-hop for those prefixes. Between the RRs in the different ASs, there is a multihop EBGP session for the exchange of VPN-IPv4 customer prefixes. The RRs do not override the next-hop for those prefixes.

The main advantage of this model is that no VPN routes need to be held on the ASBR-PEs and therefore, it scales the best among all the three Inter-AS IP-VPN models. However, leaking /32 PE addresses between service providers raises some security concerns. Therefore, we see Model C typically deployed within a service provider network.

Figure 15: Inter-AS VPN Model C shows the example topology which consists of four SR OS nodes in AS 64496 and four SR OS nodes in AS 64497. There is an AS interconnection between ASBR PE-4 to ASBR PE-8. PE-3 and PE-7 act as RRs for their AS. An IP-VPN is configured in each AS. The initial configuration includes the following:

- IS-IS or OSPF on all interfaces within each of the ASs.
- LDP on all interfaces within each of the ASs.

- MP-IBGP sessions between the PE routers and the RRs in each of the ASs, as shown in the following section.
- IP-VPN on PE-1 and on PE-5 with identical route targets.
- A loopback interface in the VRF on PE-1 and PE-5.

## Configuration

The first step is to configure an MP-IBGP session between the PEs in both ASs. An export policy is configured to export the system prefixes from the PEs in labeled BGP.

PE-3 and PE-7 act as RR in the ASs. In AS 64496, PE-1 and PE-2 are peered with RR PE-3 for the labeled IPv4 and VPN-IPv4 address families; ASBR PE-4 is peered with RR PE-3 for the labeled IPv4 address family only. In AS 64497, PE-5 and PE-6 are peered with RR PE-7 for the labeled IPv4 and VPN-IPv4 address families; ASBR PE-8 is peered with RR PE-7 for the labeled IPv4 address family only.

Address family **label-ipv4** is required to advertise labeled IPv4 routes toward each neighbor PE. Address family **vpn-ipv4** is required to advertise IPv4 customer VPN routes within the AS.

The initial BGP configuration for RR PE-3 is as follows:

```
# on RR PE-3:
configure
  router Base
    autonomous-system 64496
    bgp
      split-horizon
      group "IBGP"
        cluster 192.0.2.3
        export "export-bgp"
        peer-as 64496
        neighbor 192.0.2.1
          family vpn-ipv4 label-ipv4
          advertise-inactive
        exit
        neighbor 192.0.2.2
          family vpn-ipv4 label-ipv4
          advertise-inactive
        exit
        neighbor 192.0.2.4
          family label-ipv4
          advertise-inactive
        exit
      exit
    exit
```

The export policy is defined as follows:

```
# on PE-1, PE-2, PE-3, PE-5, PE-6, PE-7:
configure
  router Base
    policy-options
      begin
      prefix-list "PE-sys"
        prefix 192.0.2.0/28 longer
      exit
      policy-statement "export-bgp"
        entry 10
          from
```

```

        protocol direct
        prefix-list "PE-sys"
    exit
    action accept
    exit
exit
exit
commit

```

On the ASBRs in both ASs, EBGP and IBGP need to be configured. The EBGP session is configured with **advertise-inactive** and is used to redistribute labeled IPv4 routes for the /32 system IP addresses between the ASs, even if those routes are not the most preferred routes within the system for a certain destination.

The configuration for ASBR PE-4 is as follows. The address family **label-ipv4** is required to enable the advertising of labeled IPv4 routes. This address family is also required on the RR neighbor in order to propagate the labeled IPv4 routes toward the other PEs in the AS.

```

# on ASBR PE-4:
configure
router Base
    autonomous-system 64496
    bgp
        split-horizon
        group "EBGP"
            neighbor 192.168.48.2
                family label-ipv4
                peer-as 64497
                advertise-inactive
        exit
    exit
    group "IBGP"
        peer-as 64496
        neighbor 192.0.2.3
            family label-ipv4
        exit
    exit
exit
exit
exit

```

On the remaining PE nodes in AS 64496, PE-1 and PE-2, the address families **label-ipv4** and **vpn-ipv4** must be enabled, as follows:

```

# on PE-1, PE-2:
configure
router Base
    autonomous-system 64496
    bgp
        split-horizon
        group "IBGP"
            export "export-bgp"
            peer-as 64496
            neighbor 192.0.2.3
                family vpn-ipv4 label-ipv4
            exit
        exit
    exit
exit
exit
exit

```

The configuration for the nodes in AS 64497 is similar. The IP addresses can be derived from [Figure 15: Inter-AS VPN Model C](#).

The following command on ASBR PE-4 verifies that the EBGP and IBGP sessions for the labeled IPv4 address family are up:

```
*A:PE-4# 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
-----
192.0.2.3
Def. Inst          64496      11   0 00h02m36s 3/0/3 (Lbl-IPv4)
                  9       0
192.168.48.2
Def. Inst          64497      7    0 00h01m40s 3/3/3 (Lbl-IPv4)
                  8       0
-----
```

On ASBR PE-4, three inactive labeled IPv4 routes have been received from the IBGP peers and three active labeled IPv4 routes have been received via EBGP, as follows:

```
*A:PE-4# show router bgp routes label-ipv4

=====
BGP Router ID:192.0.2.4      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP LABEL-IPV4 Routes
=====
Flag  Network                LocalPref  MED
      Nexthop (Router)    Path-Id    IGP Cost
      As-Path            Label
-----
*i    192.0.2.1/32         100        None
      192.0.2.1          None        20
      No As-Path         524283
*i    192.0.2.2/32         100        None
      192.0.2.2          None        10
      No As-Path         524283
*i    192.0.2.3/32         100        None
      192.0.2.3          None        10
      No As-Path         524283
u*>i 192.0.2.5/32          None        None
      192.168.48.2       None        0
      64497               524283
u*>i 192.0.2.6/32          None        None
      192.168.48.2       None        0
      64497               524282
u*>i 192.0.2.7/32          None        None
```

```

192.168.48.2          None          0
64497                524281
-----
Routes : 6
=====

```

The following three routes have been received from EBGW peer PE-8: one for each system IP address in the remote AS, except for the ASBR itself:

```

*A:PE-4# show router bgp neighbor 192.168.48.2 received-routes label-ipv4
=====
BGP Router ID:192.0.2.4      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
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 192.0.2.5/32            n/a        None
      192.168.48.2          None        0
      64497                  524283
u*>i 192.0.2.6/32            n/a        None
      192.168.48.2          None        0
      64497                  524282
u*>i 192.0.2.7/32            n/a        None
      192.168.48.2          None        0
      64497                  524281
-----
Routes : 3
=====

```

In this example, the IP prefix for PE-8 itself is not included. The prefix of the ASBR need not be advertised in labeled BGP to the remote AS, because ASBRs do not advertise VPN-IPv4 prefixes.

More detailed information about the advertised route from PE-5 can be seen with following command on PE-4:

```

*A:PE-4# show router bgp routes 192.0.2.5/32 label-ipv4 hunt
=====
BGP Router ID:192.0.2.4      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP LABEL-IPV4 Routes
=====
RIB In Entries
-----
Network       : 192.0.2.5/32
Nexthop       : 192.168.48.2
Path Id       : None

```

```

From          : 192.168.48.2
Res. Nexthop  : 192.168.48.2
Local Pref.   : None
Aggregator AS : None
Atomic Aggr.  : Not Atomic
AIGP Metric   : None
Connector     : None
Community     : No Community Members
Cluster       : No Cluster Members
Originator Id : None
Fwd Class     : None
IPv4 Label   : 524283
Flags         : Used Valid Best IGP In-TTM In-RTM
Route Source  : External
AS-Path       : 64497
Route Tag     : 0
Neighbor-AS   : 64497
Orig Validation: NotFound
Source Class  : 0
Add Paths Send : Default
RIB Priority   : Normal
Last Modified : 00h03m58s

Interface Name : int-PE-4-PE-8
Aggregator     : None
MED            : None
IGP Cost       : 0

Peer Router Id : 192.0.2.8
Priority        : None

Dest Class     : 0
    
```

-----  
RIB Out Entries  
-----

```

Network       : 192.0.2.5/32
Nexthop       : 192.0.2.4
Path Id       : None
To            : 192.0.2.3
Res. Nexthop  : n/a
Local Pref.   : 100
Aggregator AS : None
Atomic Aggr.  : Not Atomic
AIGP Metric   : None
Connector     : None
Community     : No Community Members
Cluster       : No Cluster Members
Originator Id : None
IPv4 Label   : 524283
Lbl Allocation : NEXT-HOP
Origin        : IGP
AS-Path       : 64497
Route Tag     : 0
Neighbor-AS   : 64497
Orig Validation: NotFound
Source Class  : 0

Interface Name : NotAvailable
Aggregator     : None
MED            : None
IGP Cost       : 0

Peer Router Id : 192.0.2.3
Label Type   : SWAP

Dest Class     : 0
    
```

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

In the RIB In entries, the received label from PE-8 can be seen (524283). In the RIB Out entries, the locally assigned (Advertised) label for this prefix can be seen (524283). These labels need not match. The ASBR PE-4 swaps BGP labels, according to the following label mapping:

```
*A:PE-4# show router bgp inter-as-label
```

```

=====
BGP Inter-AS labels
Flags: B - entry has backup, P - entry is promoted
=====
NextHop                Received            Advertised          Label
    
```

	Label	Label	Origin
192.0.2.1	524283	524280	Internal
192.0.2.2	524283	524279	Internal
192.0.2.3	524283	524278	Internal
192.168.48.2	524281	524281	External
192.168.48.2	524282	524282	External
<b>192.168.48.2</b>	<b>524283</b>	<b>524283</b>	<b>External</b>
-----			
Total Labels allocated: 6			
=====			

The route from PE-1 toward PE-5 uses received label 524283 and advertised label 524283, as indicated on the sixth row in the table. The BGP label in the label stack sent by PE-1 contains BGP label 524283 toward ASBR PE-4, where it is swapped to BGP label 524283 toward ASBR PE-8.

ASBR PE-8 swaps BGP label 524283 to BGP label 524283 toward PE-5, as follows:

```
*A:PE-8# show router bgp inter-as-label
```

=====			
BGP Inter-AS labels			
Flags: B - entry has backup, P - entry is promoted			
=====			
NextHop	Received Label	Advertised Label	Label Origin
-----			
<b>192.0.2.5</b>	<b>524283</b>	<b>524283</b>	<b>Internal</b>
192.0.2.6	524283	524282	Internal
192.0.2.7	524283	524281	Internal
192.168.48.1	524278	524278	External
192.168.48.1	524279	524279	External
192.168.48.1	524280	524280	External
-----			
Total Labels allocated: 6			
=====			

On ASBR PE-4, the routes toward PE-5, PE-6, and PE-7 in the remote AS have been installed in the route table, as follows:

```
*A:PE-4# show router route-table
```

=====					
Route Table (Router: Base)					
=====					
Dest Prefix[Flags]	Type	Proto	Age	Pref	
Next Hop[Interface Name]			Metric		
-----					
192.0.2.1/32	Remote	ISIS	00h07m04s	18	
192.168.24.1			20		
192.0.2.2/32	Remote	ISIS	00h07m04s	18	
192.168.24.1			10		
192.0.2.3/32	Remote	ISIS	00h07m04s	18	
192.168.34.1			10		
192.0.2.4/32	Local	Local	00h07m05s	0	
system			0		
<b>192.0.2.5/32</b>	<b>Remote</b>	<b>BGP_LABEL</b>	<b>00h05m08s</b>	<b>170</b>	
<b>192.168.48.2</b>			<b>0</b>		
<b>192.0.2.6/32</b>	<b>Remote</b>	<b>BGP_LABEL</b>	<b>00h05m08s</b>	<b>170</b>	
<b>192.168.48.2</b>			<b>0</b>		
<b>192.0.2.7/32</b>	<b>Remote</b>	<b>BGP_LABEL</b>	<b>00h05m08s</b>	<b>170</b>	
<b>192.168.48.2</b>			<b>0</b>		

```

192.168.24.0/30          Local   Local   00h07m05s  0
    int-PE-4-PE-2          0
192.168.34.0/30          Local   Local   00h07m05s  0
    int-PE-4-PE-3          0
192.168.48.0/30          Local   Local   00h07m05s  0
    int-PE-4-PE-8          0
-----
No. of Routes: 10
Flags: n = Number of times nexthop is repeated
      B = BGP backup route available
      L = LFA nexthop available
      S = Sticky ECMP requested
=====

```

The BGP labeled routes for the remote PE system prefixes are further advertised toward all the PEs in the AS (through the RR) and are installed in the routing table on all PEs.

At this point, all PEs in one AS have the /32 system IPs of the remote PEs in their routing table, for example for PE-1:

```

*A:PE-1# show router route-table

=====
Route Table (Router: Base)
=====
Dest Prefix[Flags]          Type   Proto   Age           Pref
  Next Hop[Interface Name]  Metric
-----
192.0.2.1/32                Local  Local   00h07m25s    0
    system                    0
192.0.2.2/32                Remote  ISIS    00h07m18s    18
    192.168.12.2              10
192.0.2.3/32                Remote  ISIS    00h07m12s    18
    192.168.13.2              10
192.0.2.4/32                Remote  ISIS    00h07m05s    18
    192.168.12.2              20
192.0.2.5/32                Remote  BGP_LABEL 00h04m42s  170
    192.0.2.4 (tunneled)      20
192.0.2.6/32                Remote  BGP_LABEL 00h04m42s  170
    192.0.2.4 (tunneled)      20
192.0.2.7/32                Remote  BGP_LABEL 00h04m42s  170
    192.0.2.4 (tunneled)      20
192.168.12.0/30            Local  Local   00h07m25s    0
    int-PE-1-PE-2            0
192.168.13.0/30            Local  Local   00h07m25s    0
    int-PE-1-PE-3            0
-----
No. of Routes: 9
Flags: n = Number of times nexthop is repeated
      B = BGP backup route available
      L = LFA nexthop available
      S = Sticky ECMP requested
=====

```

All PEs in one AS have also received labels for all /32 system IP addresses of the remote PEs. Therefore, an MP-EBGP session can be created between the RRs in the different ASs to exchange VPN-IPv4 routes.

The additional BGP configuration for RR PE-3 is as follows. The configuration for RR PE-7 is similar. The IP addresses can be derived from [Figure 16: Protocol overview](#).

```

# on RR PE-3:
configure

```



```

router Base
  bgp
    group "peer-AS-RR"
      family vpn-ipv4
      peer-as 64497
      local-address 192.0.2.3
      neighbor 192.0.2.7
        multihop 10
        vpn-apply-export
      export "EBGP-VPN-IPv4"
    exit
  exit
exit

```

Policies can be applied on the peering session using the **export** command followed by a policy name, together with the **vpn-apply-export** command necessary to enforce base BGP instance policy on VPN-IPv4 prefixes.

On the RRs, the MP-EBGP session is up, as follows:

```

*A:PE-3# show router bgp neighbor 192.0.2.7

=====
BGP Neighbor
=====
-----
Peer          : 192.0.2.7
Description   : (Not Specified)
Group         : peer-AS-RR
-----
Peer AS       : 64497           Peer Port      : 179
Peer Address  : 192.0.2.7
Local AS      : 64496           Local Port     : 51192
Local Address : 192.0.2.3
Peer Type     : External       Dynamic Peer   : No
State       : Established    Last State     : Active
Last Event    : rcv0pen
Last Error    : Unrecognized Error
Local Family  : VPN-IPv4
Remote Family : VPN-IPv4
---snip---

```

The EBGP session between the two RRs is established.

The VPRNs on PE-1 in AS 64496 and PE-5 in AS 64497 are now interconnected. The route table for VPRN 1 shows that the remote PE can be reached via a BGP tunnel, as follows:

```

*A:PE-1# show router 1 route-table

=====
Route Table (Service: 1)
=====
-----
Dest Prefix[Flags]          Type   Proto   Age           Pref
  Next Hop[Interface Name]           Metric
-----
10.1.1.1/32                 Local  Local   00h09m51s    0
  loopback                   0
10.2.2.2/32                 Remote BGP VPN 00h00m29s    170
  192.0.2.5 (tunneled:BGP)         1000
-----
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
=====
```

Packets originating in AS 64496 with a destination in AS 64497 have 3 labels in AS 64496 (and in AS 64497). Originate a VPRN ping on PE-1 toward the VPRN loopback IP address on PE-5:

```
*A:PE-1# ping router 1 10.2.2.2
PING 10.2.2.2 56 data bytes
64 bytes from 10.2.2.2: icmp_seq=1 ttl=64 time=6.25ms.
64 bytes from 10.2.2.2: icmp_seq=2 ttl=64 time=6.39ms.
64 bytes from 10.2.2.2: icmp_seq=3 ttl=64 time=5.90ms.
64 bytes from 10.2.2.2: icmp_seq=4 ttl=64 time=6.31ms.
64 bytes from 10.2.2.2: icmp_seq=5 ttl=64 time=6.85ms.

---- 10.2.2.2 PING Statistics ----
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min = 5.90ms, avg = 6.34ms, max = 6.85ms, stddev = 0.304ms
```

The top label is the LDP label to reach the exit point of the AS (PE-4). This label has value 524284, as can be seen with following command on PE-1:

```
*A:PE-1# show router ldp bindings active prefixes prefix 192.0.2.4/32
=====
LDP Bindings (IPv4 LSR ID 192.0.2.1)
(IPv6 LSR ID ::)
=====
Label Status:
  U - Label In Use, N - Label Not In Use, W - Label Withdrawn
  WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
  e - Label ELC
FEC Flags:
  LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
  BA - ASBR Backup FEC
  (S) - Static          (M) - Multi-homed Secondary Support
  (B) - BGP Next Hop    (BU) - Alternate Next-hop for Fast Re-Route
  (I) - SR-ISIS Next Hop (O) - SR-OSPF Next Hop
  (C) - FEC resolved with class-based-forwarding
=====
LDP IPv4 Prefix Bindings (Active)
=====
Prefix          Op
IngLbl          EgrLbl
EgrNextHop      EgrIf/LspId
-----
192.0.2.4/32    Push
--             524284
192.168.12.2    1/1/1

192.0.2.4/32    Swap
524284         524284
192.168.12.2    1/1/1

-----
No. of IPv4 Prefix Active Bindings: 2
=====
```

This LDP label is popped by ASBR PE-4. No LDP label is used between the ASBRs. ASBR PE-8 pushes another LDP label.

To reach a PE in the remote AS, a BGP transport label is required, which is the middle label in the stack. The tunnel table on PE-1 shows a BGP tunnel toward PE-5, as follows:

```
*A:PE-1# show router tunnel-table

=====
IPv4 Tunnel Table (Router: Base)
=====
Destination          Owner      Encap TunnelId Pref  Nexthop      Metric
  Color
-----
192.0.2.2/32         ldap      MPLS  65537    9    192.168.12.2  10
192.0.2.3/32         ldap      MPLS  65538    9    192.168.13.2  10
192.0.2.4/32         ldap      MPLS  65539    9    192.168.12.2  20
192.0.2.5/32       bgp      MPLS  262145  12  192.0.2.4  1000
192.0.2.6/32         bgp       MPLS  262146   12   192.0.2.4    1000
192.0.2.7/32         bgp       MPLS  262147   12   192.0.2.4    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 BGP label is assigned by the next hop, in this case by the local ASBR PE-4. This IPv4 label can be seen with following command on PE-1:

```
*A:PE-1# show router bgp routes 192.0.2.5/32 label-ipv4 hunt

=====
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
Origin codes : i - IGP, e - EGP, ? - incomplete

=====
BGP LABEL-IPV4 Routes
=====
RIB In Entries
-----
Network       : 192.0.2.5/32
Nexthop       : 192.0.2.4
Path Id       : None
From          : 192.0.2.3
Res. Nexthop  : 192.0.2.4 (LDP)
Local Pref.   : 100
Aggregator AS : None
Atomic Aggr.  : Not Atomic
AIGP Metric   : None
Connector     : None
Community     : No Community Members
Cluster       : 192.0.2.3
Originator Id : 192.0.2.4
Fwd Class     : None
IPv4 Label   : 524283
Flags         : Used Valid Best IGP In-TTM In-RTM
Route Source  : Internal
AS-Path       : 64497
Route Tag     : 0
Neighbor-AS   : 64497
Interface Name : NotAvailable
Aggregator    : None
MED           : None
IGP Cost      : 20
Peer Router Id : 192.0.2.3
Priority       : None
```

```

Orig Validation: NotFound
Source Class   : 0                      Dest Class   : 0
Add Paths Send : Default
RIB Priority   : Normal
Last Modified  : 00h10m56s

-----
RIB Out Entries
-----
-----
Routes : 1
=====

```

This BGP label is swapped by ASBR PE-4 in AS 64496 and by ASBR PE-8 in AS 64497.

The bottom label is the VPN label assigned by the remote PE in the remote AS for the destination network. This VPN label is retrieved on PE-1, as follows:

```

*A:PE-1# show router bgp routes 10.2.2.2/32 vpn-ipv4 hunt
=====
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
Origin codes  : i - IGP, e - EGP, ? - incomplete

=====
BGP VPN-IPv4 Routes
=====
-----
RIB In Entries
-----
Network       : 10.2.2.2/32
Nexthop       : 192.0.2.5
Route Dist.   : 64497:1          VPN Label    : 524282
Path Id       : None
From          : 192.0.2.3
Res. Nexthop  : n/a
Local Pref.   : 100
Aggregator AS : None            Interface Name : NotAvailable
Atomic Aggr.  : Not Atomic      Aggregator    : None
AIGP Metric   : None            MED           : None
Connector     : None            IGP Cost      : 0
Community     : target:64497:1
Cluster       : No Cluster Members
Originator Id : None            Peer Router Id : 192.0.2.3
Fwd Class     : None            Priority       : None
Flags         : Used Valid Best IGP
Route Source  : Internal
AS-Path       : 64497
Route Tag     : 0
Neighbor-AS   : 64497
Orig Validation: N/A
Source Class  : 0                      Dest Class   : 0
Add Paths Send : Default
Last Modified  : 00h02m23s
VPRN Imported  : 1

-----
RIB Out Entries
-----
-----

```

```
Routes : 1
```

## Conclusion

Inter-AS option C allows the delivery of Layer 3 VPN services to customers who have sites connected in different ASs. This example shows the configuration of inter-AS option C (specific to this feature) together with the associated show output which can be used for verification and troubleshooting.

# Intra-AS NG-MVPN over BIER

This chapter provides information about Intra-AS NG-MVPN over BIER.

Topics in this chapter include:

- [Applicability](#)
- [Overview](#)
- [Configuration](#)
- [Conclusion](#)

## Applicability

The information and configuration in this chapter are based on SR OS Release 16.0.R7.

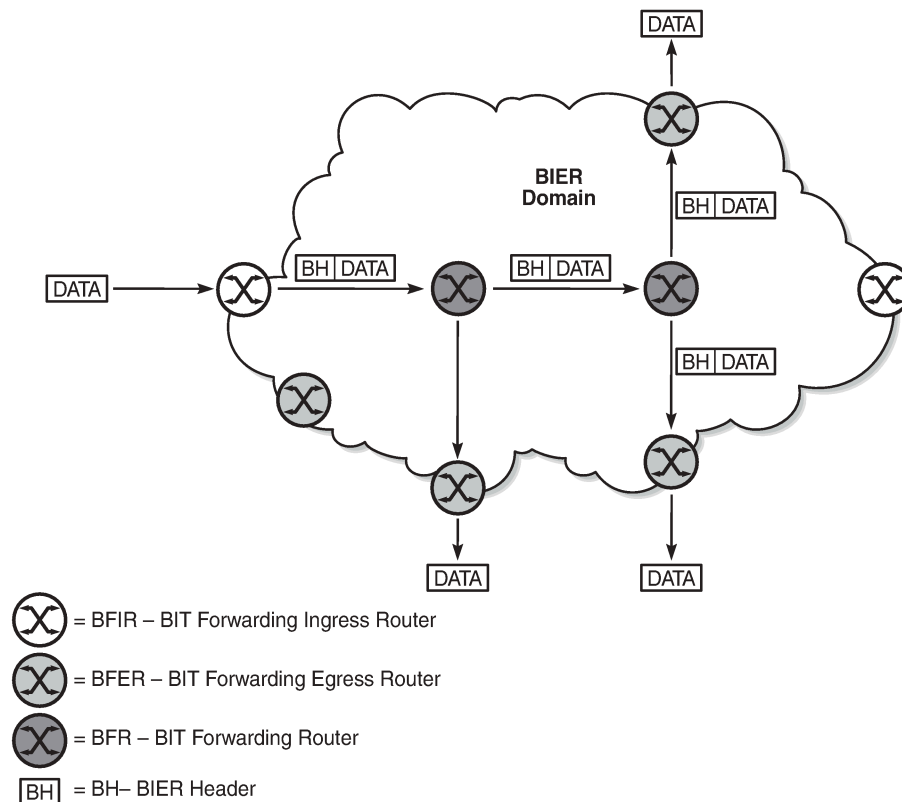
## Overview

RFC 8279, *Multicast Using Bit Index Explicit Replication (BIER)*, defines an architecture for the forwarding of multicast data packets through a multicast domain, without requiring any explicit tree-building protocol nor any intermediate nodes to maintain any per-flow state, thereby offering significant operational simplification.

## BIER concepts

BIER-enabled routers are known as Bit Forwarding Routers (BFRs). A BIER domain contains Bit Forwarding Ingress Routers (BFIRs), Bit Forwarding Egress Routers (BFERs), and transit BFRs; see [Figure 17: Bit Forwarding Router types](#). A router can be a BFIR for one flow, and at the same time be a BFER or a transit BFR for other flows. A BFIR adds a BIER header holding the information used by the BIER forwarding procedures to the multicast packets entering the BIER domain. A BFER removes the BIER header when forwarding the packets out of the BIER domain. The BIER encapsulated data can be further encapsulated in MPLS, where a BIER forwarding table is identified through an MPLS label on the adjacent node. BIER tunnels can build a fully meshed multicast network, thereby providing multicast interconnections between all the PEs in the network.

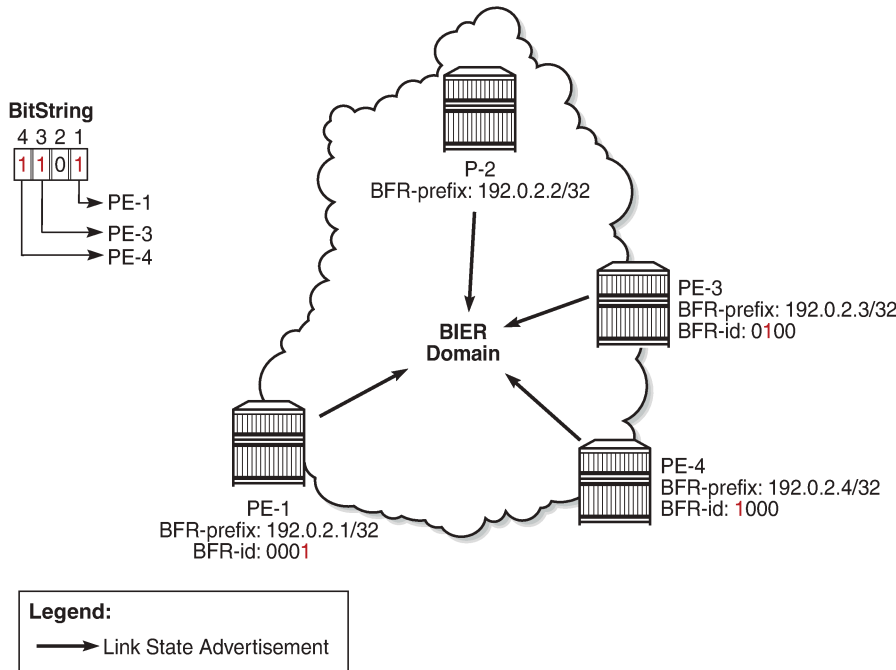
Figure 17: Bit Forwarding Router types



28832

Every BIER receiver has a BFR-prefix and a BFR-id. The BFR-id identifies a unique bit position in a bit string assigned to that BFR; by convention, the rightmost bit is bit number 1. The mapping between the BFR-id and the BFR-prefix must be known to all BFRs in the domain; therefore, this information is distributed by the underlying Interior Gateway Protocol (IGP) in new TLVs and sub-TLVs defined in IGP extensions. In the example in [Figure 18: BIER control plane: example bit position assignment and advertisement](#), PE-1 has BFR-prefix 192.0.2.1/32 and BFR-id 1, PE-3 has BFR-prefix 192.0.2.3/32 and BFR-id 3, and so on. Routers in the core of the network are transit BFRs and, as such, they are not BIER receivers; they require a BFR-prefix but not a BFR-id, so their BFR-id is zero.

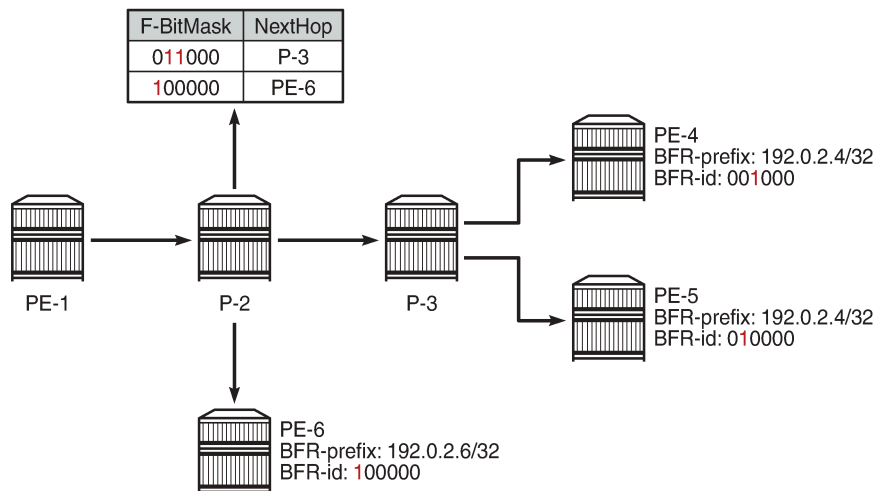
Figure 18: BIER control plane: example bit position assignment and advertisement



28833

Every BFR in the domain constructs a Bit Indexed Forwarding Table (BIFT) using the shortest unicast path route (SPF). In the construction of a BIFT, every BFR computes the unicast SPF path to each BFR-prefix. BFR-ids sharing the same next hop are combined using a logical OR operation, thereby saving memory resources by occupying a single entry in the BIFT; see [Figure 19: BIER data plane: example BIER forwarding table for P-2](#) for an example. On P-2, PE-4 and PE-5 are reachable via P-3, so the forwarding bitmask is 011000 with next hop P-3.

Figure 19: BIER data plane: example BIER forwarding table for P-2



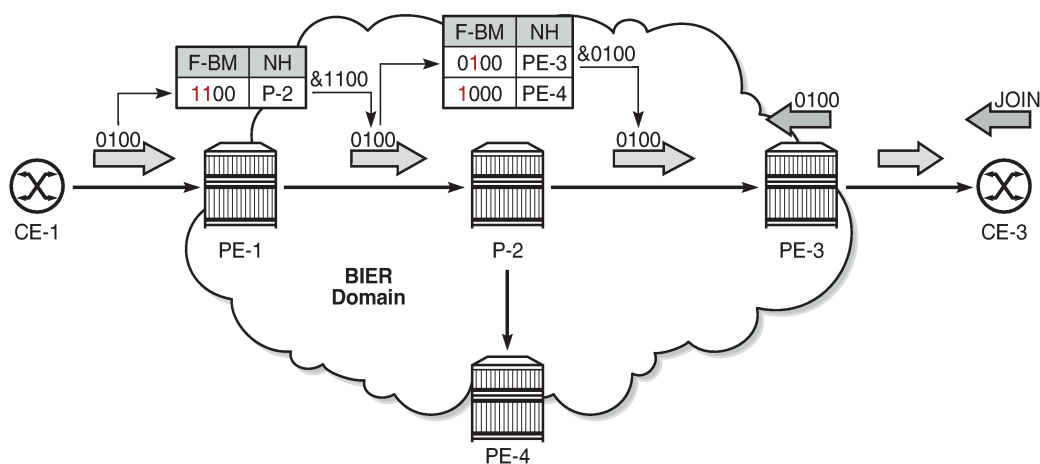
28834



The set of receivers a multicast data packet must be sent to is encoded in a bit string that is embedded in the BIER header. A BFR receiving a multicast data packet uses its BIFT to replicate these packets. When replicating a packet, the bit string in the BIER header is rewritten by the BFRs to avoid loops. Packets are forwarded and replicated hop by hop, following the unicast path from the BFIR to the BFERs.

The example in [Figure 20: BIER data plane: example BIER packet forwarding](#) assumes that CE-3 sends an IGMP join to PE-3, which in turn signals the join in MP-BGP to PE-1. PE-1 searches and finds BFR-id 3 in its BIFT with P-2 as the next hop, and logically ANDs (&1100) its bitmap when forwarding the packet to P-2. The AND operation explicitly clears bits for destinations that do not need the packet, thereby preventing potential duplication, and avoiding multicast routing loops. P-2 performs a similar set of steps to forward the packet to PE-3, which delivers the packet to CE-3. The overall result is that only nodes that requested the multicast stream will get that stream, leading to better, more optimal network usage.

Figure 20: BIER data plane: example BIER packet forwarding



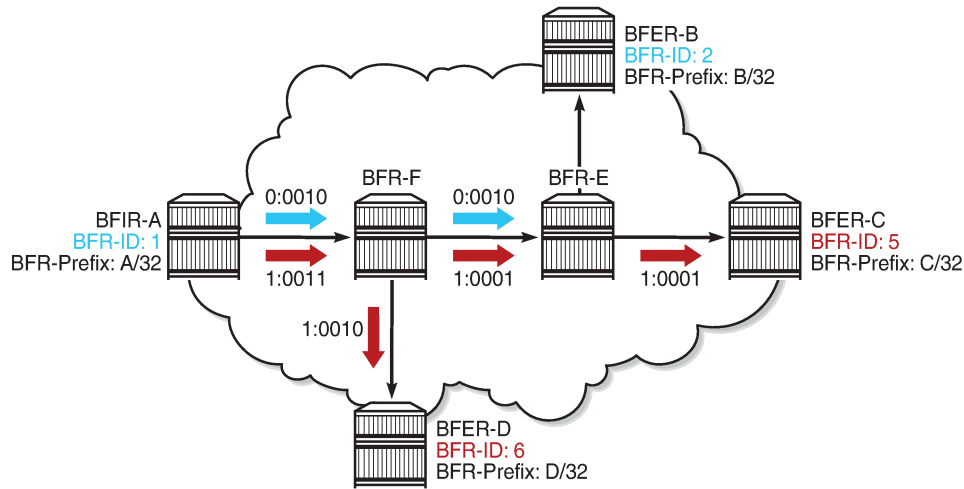
28835

A BIER domain contains one or more sub-domains (SDs), where each SD is associated with a single IS-IS or OSPF topology. Each SD is identified by a number in the range 0 to 255, and each BIER domain must contain at least one SD, where SD 0 is the default. If a BIER domain contains more than one SD, each BFR in the domain must be provisioned with the set of SDs it belongs to. The BFR-id of a BFR is a number in the range 1 to 65535, and must be unique within an SD. If a BFR belongs to more than one SD, it may have different BFR-ids in each SD. Nokia recommends using a loopback interface for the BFR-prefix. The BFR-prefix to BFR-id mapping is flooded within the SD.

For scalability, a BIER domain contains one or more BIER sets, where each BIER set is identified by a Set Identifier [SI]. The Bit String Length (BSL) dictates how many BFRs can be represented in a BIER set. The BSL, the SI, and the Bit Position (BP) are encoded in the BIER header, where the SI and the BP are derived from the BFR-id. Assuming a BSL of 256, BPs can range from 1 to 256 in BIER set 0. If more than 256 BFIRs and BFERs are required, a second BIER set is required with SI 1, where the BP can again range from 1 to 256. However, if a multicast flow has multiple receivers in different BIER sets on the same outgoing interface, the packet must be replicated to every BIER set.

[Figure 21: BIER sets](#) provides an example where BSL is 4 bits, and BFER B, C, and D are interested in the same stream entering the BIER domain at BFIR A. Because BFER C and D have BFR-id 5 and 6, respectively, they are part of SI 1, and BFIR A has to make two copies of the stream; a first (blue) copy to reach BFER-B, and a second (red) copy to reach BFER C and BFER D. Therefore, Nokia recommends assigning BFR-ids as dense as possible; for example, in consecutive order starting from 1.

Figure 21: BIER sets



28836

BIER is encapsulated in MPLS and the MPLS label for each forwarding table (identified through BSL, SI, and SD) is distributed through IS-IS; the BSL is 256 and the maximum number of BIER sets is 16. RFC 8401, *Bit Index Explicit Replication Support via IS-IS*, defines the extensions needed for distributing BIER information in IS-IS.

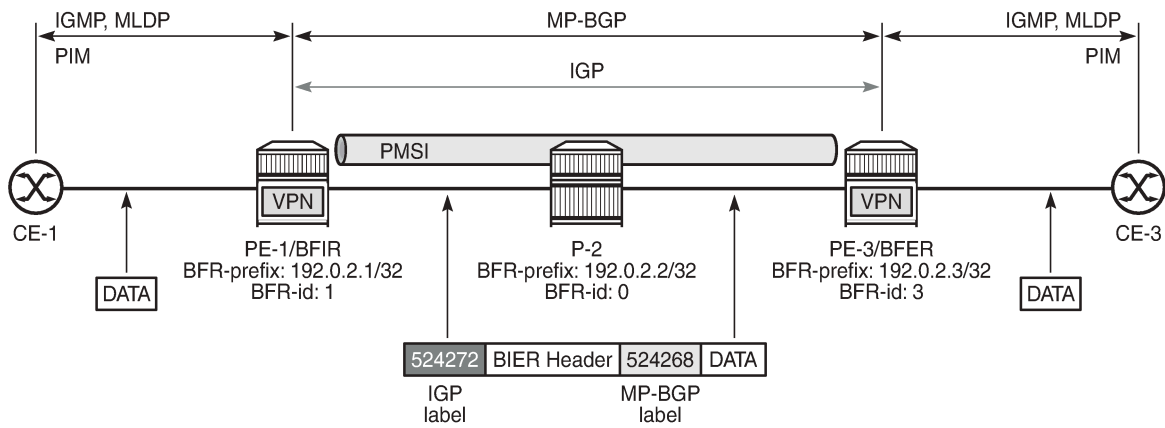
### Multicast VPN over BIER

Multicast VPN (MVPN) or Next Generation IP Multicast in an IP-VPN (NG-MVPN) architectures describe a set of virtual routing and forwarding (VRF) or virtual private routed networks (VPRNs) that support the transport of multicast traffic across a provider network. MVPNs are defined in RFC 6513, *Multicast in MPLS/BGP IP VPNs*, and RFC 6514, *BGP Encodings and Procedures for Multicast in MPLS/IP VPNs*.

The *NG-MVPN Configuration with MPLS* and *NG-MVPN Configuration with PIM* chapters provide examples where the provider tunnels are signaled through either mLDP or PIM.

In this chapter, MVPN is used as the overlay to BIER in a single Autonomous System (AS). The MP-BGP control plane is used for the Auto-Discovery (A-D) of the MVPN memberships, the provider tunnel (P-tunnel) signaling, and the customer route (C-route) signaling; see [Figure 22: MVPN over BIER](#).

Figure 22: MVPN over BIER



28837

BIER supports the use of Inclusive PMSIs (I-PMSIs) and Selective PMSIs (S-PMSIs). In the case of I-PMSI, the A-D route signaling, P-tunnel instantiation, and C-multicast routing information are restricted to the I-PMSI, where every BFIR and BFER participating in the MVPN receives every packet forwarded onto the I-PMSI. This way BFIRs and BFERs are interconnected in full mesh. Bandwidth utilization can be optimized by using S-PMSIs, where the C-flow is sent only to BFERs that have interested receivers, and where explicit tracking is used to create the list of interested receivers. The C-flow will be moved from I-PMSI to S-PMSI by the BFIR when a configured data threshold is reached.

The MP-BGP UPDATE messages used to establish the I-PMSI and the S-PMSI tunnels include the PMSI Tunnel Attribute (PTA), where the tunnel type is set to BIER, and where the MPLS label value is the upstream assigned MPLS label that identifies the VRF; see [Figure 23: PTA: PMSI tunnel attribute](#). By using BIER, multicast provisioning in the core can be simplified; the core routers must be provisioned for BIER, but not for PIM, mLDP, or RSVP-TE. A PIM-free core can be created where no multicast state needs to be maintained.

Figure 23: PTA: PMSI tunnel attribute

PMSI Tunnel Attribute (PTA)		Field	Purpose
Tunnel Identifier	Flags (1 octet)	Flags	LIR-bit. Leaf-Information Required. Used for Explicit Tracking
	Tunnel Type = 0x0B (1 octet)	Tunnel Type	PMSI Tunnel Type; 0x0B for BIER
	MPLS Label (3 octets)	MPLS Label	Upstream-assigned non-zero MPLS Label
	Subdomain ID (1 octet)	Subdomain ID	BIER subdomain ID
	BFR-id (2 octets)	BFR-id	BFR-id of the BFIR constructing the PMSI Tunnel Attribute (PTA)
	BFR-prefix (4 or 16 octets)	BFR-prefix	BFR-prefix of the BFIR that is constructing the PTA

28838

A multicast client joining a group through IGMP, MLD, or PIM results in the VRF sending an MP-BGP Source-Join message to the BFIR. The BFIR responds with a Source-AD message and establishes an S-PMSI tunnel if S-PMSI is enabled for the VPRN. Next, the S-PMSI tunnel is used to transport the C-flow from the BFIR to the BFER. Otherwise, the I-PMSI tunnel is used for transporting the C-flow.

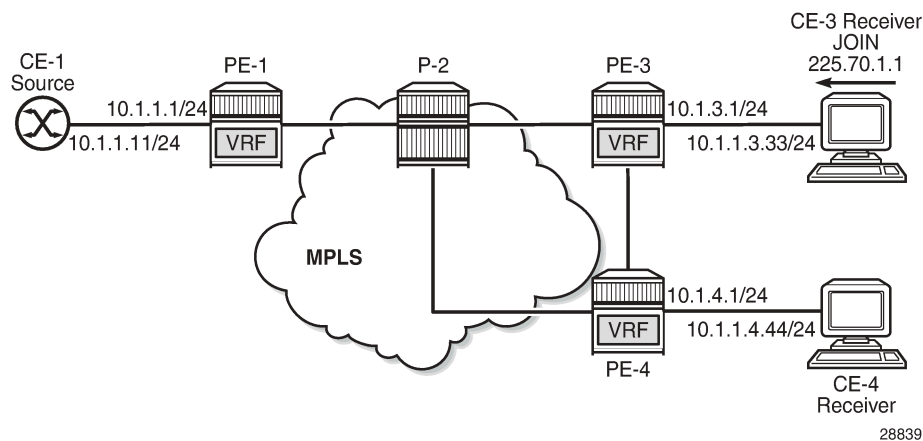
On CE to PE links, PIM Hello messages are used to establish C-PIM adjacencies. Also between the PEs, across the I-PMSIs, adjacencies are established, not using PIM but using MP-BGP instead. The announcement of MP-REACH-NLRI Intra-AS I-PMSI A-D routes in the discovery process serves as the

means to establish the PE-PE adjacencies. A response of the corresponding MP-UNREACH-NLRI results in the adjacency being dropped.

## Configuration

The configuration used in this chapter is shown in [Figure 24: Intra-AS NG-MVPN over BIER](#). IS-IS is the interior gateway protocol (IGP) used in AS 64496, and all nodes are at IS-IS level 2. These nodes share the same BIER domain and sub-domain (0). A multicast stream with group address 225.70.1.1 is transmitted by source 10.1.1.11 connected to PE-1. CE-3 and CE-4 are multicast receivers connected to PE-3 and PE-4, respectively. VPRN 1 is defined in PE-1, PE-3, and PE-4, and uses NG-MVPN for transporting the multicast traffic through the core of the network. See the [NG-MVPN Configuration with MPLS](#) and the [NG-MVPN Configuration with PIM](#) chapters for more information about NG-MVPN. The BFR-ids used by PE-1, PE-3, and PE-4 are 1, 3, and 4, respectively, so they are part of a single set with SI 0.

Figure 24: Intra-AS NG-MVPN over BIER



The initial configuration on the PE nodes includes the following:

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

## BGP configuration

BGP is required at the core of the network, using the VPN IPv4 and MVPN IPv4 address families in support of unicast and multicast for VPRN services. PE-1, PE-3, and PE-4 are clients of the route reflector located in P-2. The BGP configuration on PE-1, PE-3, and PE-4 is as follows:

```
# on PE-1, PE-3, and PE-4
configure
router
  autonomous-system 64496
  bgp
    family vpn-ipv4 mvpn-ipv4
```

```

        vpn-apply-import
        vpn-apply-export
        rapid-withdrawal
        rapid-update vpn-ipv4 mvpn-ipv4
        group "iBGP"
            neighbor 192.0.2.2
                peer-as 64496
            exit
        exit
        no shutdown
    exit
exit
exit
exit

```

P-2 is the route reflector and configured as follows:

```

# on P-2
configure
router
    autonomous-system 64496
    bgp
        family vpn-ipv4 mvpn-ipv4
        vpn-apply-import
        vpn-apply-export
        cluster 0.0.0.1
        rapid-withdrawal
        rapid-update vpn-ipv4 mvpn-ipv4
        group "iBGP"
            peer-as 64496
            neighbor 192.0.2.1
                exit
            neighbor 192.0.2.3
                exit
            neighbor 192.0.2.4
                exit
        exit
        no shutdown
    exit
exit
exit
exit

```

## BIER configuration

All the nodes in the topology have a BIER template named *bier-templ0*; however, the content for that template is different on each node. Although multiple sub-domains can be defined through the range command, in this example all nodes in the topology are in the single sub-domain 0 of the BIER domain.

PE-1, PE-3, and PE-4 are the termination points for the BIER tunnels; therefore, they require a BFR-id. For PE-1, PE-3, and PE-4, the system addresses are used as the prefix, and the BFR-ids are set to 1, 3, and 4, respectively. For brevity, only the BIER template on PE-1 is shown:

```

# on PE-1
configure
router
    bier
        template "bier-templ0"
            sub-domain 0
                prefix 192.0.2.1
                bfr-id 1

```

```

        exit
        no shutdown
    exit
    no shutdown
exit
exit
exit

```

P-2 is a transit BFR, so P-2 does not require a BFR-id. This is made explicit in the BIER template for P-2, as follows:

```

# on P-2
configure
router
    bier
        template "bier-templ0"
            sub-domain 0
            prefix 192.0.2.2
            no bfr-id
        exit
        no shutdown
    exit
    no shutdown
exit
exit
exit

```

The BIER template must be applied to the IGP, so that the MPLS labels required for the BIER tunnels are distributed through the IGP, which is IS-IS.

In the example topology shown in Figure 8, the IS-IS configuration is similar on all nodes, with **level-capability** set to **level-2**, and *bier-templ0* applied and enabled (no shutdown) at level 2, as follows. For brevity, only the IS-IS configuration for PE-1 is shown.

```

# on PE-1
configure
router
    isis
        level-capability level-2
        area-id 49.0001
        traffic-engineering
        level 2
            bier
                template "bier-templ0"
                no shutdown
            exit
        exit
        interface "system"
            no shutdown
        exit
        interface "int-PE-1-P-2"
            interface-type point-to-point
            no shutdown
        exit
        no shutdown
    exit
exit
exit
exit

```

The status of IS-IS shows that BIER is active at L2 using template bier-tmpl0, but not at L1. BIER must be active on all nodes in the topology, but for brevity only the IS-IS status on P-1 is shown, as follows:

```
*A:PE-1# show router isis status

=====
Rtr Base ISIS Instance 0 Status
=====
ISIS Cfg System Id       : 0000.0000.0000
ISIS Oper System Id     : 1920.0000.2001
ISIS Cfg Router Id      : 0.0.0.0
ISIS Oper Router Id     : 192.0.2.1
ASN                      : 0
Admin State              : Up
Oper State                : Up
Ipv4 Routing             : Enabled
---snip---
L1 Bier Template       : None, Disabled
L2 Bier Template       : bier-tmpl0, Enabled
---snip---
Entropy Label            : Enabled
Override ELC              : Disabled
=====
*A:PE-1#
```

BIER allocates an MPLS label per BIER set, and 16 consecutive labels are allocated to accommodate the 16 BIER sets supported by SR OS. This label range is shown together with the BFR-id and BFR-prefix from the BIER template in the BIER database. For PE-1, the BIER database is as follows:

```
*A:PE-1# show router bier database

=====
BIER Database
=====
Template          MT          Sub-domain      BSL
BFR-ID          Start         End           Total
BFR-Prefix
-----
bier-tmpl0        ipv4-unicast    0                256
1                524271         524286        16
192.0.2.1
=====
BIER Database entries : 1
=====
*A:PE-1#
```

On P-2, the BIER database looks similar, but the BFR-id is zero indicating that no BFR-id is defined, as follows:

```
*A:P-2# show router bier database

=====
BIER Database
=====
Template          MT          Sub-domain      BSL
BFR-ID          Start         End           Total
BFR-Prefix
-----
bier-tmpl0        ipv4-unicast    0                256
0                524272         524287        16
=====
*A:P-2#
```

```
192.0.2.2
```

```
=====
BIER Database entries : 1
=====
```

```
*A:P-2#
```

IS-IS distributes the BIER information across all the BFRs in the network using link state packets (shown as "LSP" in the output); consequently, this information is the same on all BFRs. As an example, the details for LSP PE-3.00-00 show that for PE-3 at level 2, the BFR-id is 3, and the MPLS label value is 524271:

```
*A:PE-1# show router isis database PE-3.00-00 detail
```

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

```
Displaying Level 1 database
```

```
-----
Level (1) LSP Count : 0
```

```
Displaying Level 2 database
```

```
-----
LSP ID   : PE-3.00-00           Level   : L2
Sequence : 0x9                  Checksum : 0x5e35   Lifetime : 962
Version  : 1                    Pkt Type : 20      Pkt Ver  : 1
Attributes: L1L2                Max Area : 3       Alloc Len : 221
SYS ID   : 1920.0000.2003       SysID Len : 6      Used Len  : 221
```

```
TLVs :
```

```
Area Addresses:
  Area Address : (3) 49.0001
Supp Protocols:
  Protocols    : IPv4
IS-Hostname    : PE-3
Router ID      :
  Router ID    : 192.0.2.3
```

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

```
TE IP Reach :
  Default Metric : 10
  Control Info:  , prefLen 30
  Prefix : 192.168.23.0
  Default Metric : 10
  Control Info:  , prefLen 30
  Prefix : 192.168.34.0
  Default Metric : 0
  Control Info: S, prefLen 32
  Prefix : 192.0.2.3
Sub TLV :
  Bier::Bier Algo:0, IGP Algo:0, SD id:0, BFR id:3
  MPLS Encap Max SI:16, BS Len:3(256), Label:524271
```

```
Level (2) LSP Count : 1
```

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

```
=====
*A:PE-1#
```



A shorter and more convenient way for obtaining the BIER information directly is through the **bier-info** command. On PE-3, the BIER information is as follows:

```
*A:PE-3# show router isis bier-info

=====
Rtr Base ISIS Instance 0 Bier Info
=====

Displaying Level 1 BIER info
-----

Displaying Level 2 BIER info
-----
LSP ID   : PE-1.00-00
MT ID    : 0
Prefix   : 192.0.2.1
Sub TLV   :
    Bier::Bier Algo:0, IGP Algo:0, SD id:0, BFR id:1
    MPLS Encap Max SI:16, BS Len:3(256), Label:524271

LSP ID   : P-2.00-00
MT ID    : 0
Prefix   : 192.0.2.2
Sub TLV   :
    Bier::Bier Algo:0, IGP Algo:0, SD id:0, BFR id:0
    MPLS Encap Max SI:16, BS Len:3(256), Label:524272

LSP ID   : PE-3.00-00
MT ID    : 0
Prefix   : 192.0.2.3
Sub TLV   :
    Bier::Bier Algo:0, IGP Algo:0, SD id:0, BFR id:3
    MPLS Encap Max SI:16, BS Len:3(256), Label:524271

LSP ID   : PE-4.00-00
MT ID    : 0
Prefix   : 192.0.2.4
Sub TLV   :
    Bier::Bier Algo:0, IGP Algo:0, SD id:0, BFR id:4
    MPLS Encap Max SI:16, BS Len:3(256), Label:524271

=====
*A:PE-3#
```

No BIER tunnels are available, because no MVPN-enabled services are created yet. As stated before, SR OS allocates one label per BIER set, and on PE-1 they are as follows:

```
*A:PE-1# show router mpls-labels label 32 524287 bier

=====
MPLS Labels from 32 to 524287 (Owner: BIER)
=====
Label           Label Type      Label Owner
-----
524271          dynamic        BIER
524272          dynamic        BIER
---snip---
524285          dynamic        BIER
524286          dynamic        BIER
-----
In-use labels (Owner: BIER) in specified range : 16
```

```
In-use labels (Owner: All) in specified range : 17
In-use labels in entire range : 17
=====
*A:PE-1#
```

Based on the link state packets distributed across the network, every BFR generates a BIER routing table and a BIER forwarding table. The BIER routing table defines the interface, next-hop, and neighbor to use for all BFRs. On P-1, the BIER routing table is as follows:

```
*A:PE-1# show router bier routing
=====
Destination Prefix          Bfr-ID    Age
Neighbor
Nexthop
Interface
-----
=====
BIER Routing Database Sub-Domain 0 BSL 256
=====
192.0.2.2                   0         0d 00:27:30
192.0.2.2
192.168.12.2
int-PE-1-P-2
192.0.2.3                   3         0d 00:27:30
192.0.2.2
192.168.12.2
int-PE-1-P-2
192.0.2.4                   4         0d 00:27:30
192.0.2.2
192.168.12.2
int-PE-1-P-2
=====
Total (Sub-Domain 0): 3
=====
Total BIER Routing entries : 3
=====
*A:PE-1#
```

The BIER forwarding table has one entry per BIER neighbor, defining the next hop, interface, and forwarding bit mask. On PE-1, the BIER forwarding table is as follows. For neighbor 192.0.2.2, the forwarding bit mask is 0xC in hexadecimal, or 0b1100 in binary. Using the convention that the rightmost bit is bit 1, this means bits 3 and 4 are set, and these bits correspond to BFRs with system addresses 192.0.2.3 and 192.0.2.4, according their BFR-id.

```
*A:PE-1# show router bier forwarding
=====
Neighbor
Nexthop
Interface
[SI]: Label
Forwarding Bit Mask
BFR-ID : Prefix
-----
=====
```



```

                no shutdown
                exit
                data-threshold 224.0.0.0/4 10
            exit
        exit
        vrf-target unicast
        exit
    exit
    no shutdown
exit
exit
exit

```

The MVPN configuration for VRF 1 on PE-3 and PE-4 is the same. However, PE-3 and PE-4 provide connections to multicast clients CE-3 and CE4, and they have IGMP configured. Because the configurations for PE-3 and PE-4 are similar, only the configuration of PE-3 is provided.

```

# on PE-3
configure
service
  vprn 1
    route-distinguisher 64496:1
    auto-bind-tunnel
    resolution any
    exit
    vrf-target target:64496:1
    interface "int-PE-3-CE-3" create
      address 10.1.3.1/24
      sap 1/1/c5/1:10 create
    exit
  exit
  igmp
    ssm-translate
      grp-range 225.70.1.1 225.70.255.255
      source 10.1.1.11
    exit
  exit
  interface "int-PE-3-CE-3"
    no shutdown
  exit
  no shutdown
  exit
  pim
    no shutdown
  exit
  ---snip---
exit
exit

```

With VRF 1 on PE-1, PE-3, and PE-4 configured as previously described, VPN and MVPN routes are exchanged, as follows:

```
*A:PE-1# 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
-----
192.0.2.2
Def. Instance  64496      59   0 00h22m31s 3/2/1 (VpnIPv4)
                53   0          3/2/1 (MvpnIPv4)
-----
*A:PE-1#

```

The MVPN status for VRF-1 can be verified, as follows. For brevity, only the status on PE-1 is shown; BIER is used for I-PMSI and S-PMSI in SD 0, and the I-PMSI tunnel name is *mpls-if-73731*.

```

*A:PE-1# show router 1 mvpn

=====
MVPN 1 configuration data
=====
signaling          : Bgp                auto-discovery      : Default
UMH Selection      : Highest-Ip             SA withdrawn        : Disabled
intersite-shared   : Enabled                 Persist SA          : Disabled
vrf-import         : N/A
vrf-export         : N/A
vrf-target         : unicast
C-Mcast Import RT : target:192.0.2.1:2

ipmsi             : bier
sub-domain       : 0
i-pmsi P2MP AdmSt : Up
i-pmsi Tunnel Name : mpls-if-73731

BSR signalling     : none
Wildcard s-pmsi   : Disabled
Multistream-SPMSI : Disabled
spmsi            : bier
sub-domain       : 0
s-pmsi P2MP AdmSt : Up
max-p2mp-spmsi    : 10
data-delay-interval : 3 seconds
enable-asm-mdt    : N/A
data-threshold     : 224.0.0.0/4 --> 10 kbps

=====
*A:PE-1#

```

The receiver located at address 10.1.3.33/24 and connected to PE-3 then joins group 225.70.1.1, so VRF 1 creates an IGMP state, and (\*,225.70.1.1) is forwarded to interface *int-PE-3-CE-3*, as follows:

```

*A:PE-3# show router 1 igmp group interfaces

=====
IGMP Interface Groups
=====

(*,225.70.1.1)                               UpTime: 0d 00:10:59
  Fwd List  : int-PE-3-CE-3

-----
Entries : 1
=====
*A:PE-3#

```

VRF 1 on PE-3 also creates a PIM state for group 225.70.1.1, where the incoming S-PMSI interface is *mpls-if-73734*, and the outgoing interface is *int-PE-3-CE-3*, as follows:

```
*A:PE-3# show router 1 pim group detail

=====
PIM Source Group ipv4
=====
Group Address      : 225.70.1.1
Source Address      : 10.1.1.11
RP Address          : 0
Advt Router         : 192.0.2.1
Flags               :                               Type           : (S,G)
Mode                : sparse
MRIB Next Hop      : 192.0.2.1
MRIB Src Flags     : remote
Keepalive Timer Exp: 0d 00:02:29
Up Time            : 0d 00:12:13           Resolved By         : rtable-u

Up JP State        : Joined                Up JP Expiry        : 0d 00:00:25
Up JP Rpt         : Not Joined StarG      Up JP Rpt Override  : 0d 00:00:00

Register State    : No Info
Reg From Anycast RP: No

Rpf Neighbor      : 192.0.2.1
Incoming Intf     : mpls-if-73732
Incoming SPMSI Intf: mpls-if-73734
Outgoing Intf List : int-PE-3-CE-3

Curr Fwding Rate  : 59.920 kbps
Forwarded Packets : 3772                Discarded Packets  : 0
Forwarded Octets  : 5650456             RPF Mismatches     : 0
Spt threshold     : 0 kbps                ECMP opt threshold : 7
Admin bandwidth   : 1 kbps

-----
Groups : 1
=====
*A:PE-3#
```

The number used in the incoming S-PMSI interface from the previous command is the incoming BIER tunnel ID on PE-3. The properties of BIER tunnel 73734 indicate that this tunnel originates on BFR 1 with prefix 192.0.2.1 and uses MPLS label 524269, as follows:

```
*A:PE-3# show router bier tunnel tunnel-id 73734

=====
BIER Tunnels
=====
Tunnel-id      Type      Oper      No. Of Leaves
BFR Prefix    Bfr-ID   Mpls Label Sub-domain
-----
73734        rx      In service 0
192.0.2.1    1      524269   0

=====
BIER Tunnel entries : 1
=====
*A:PE-3#
```

The properties of the S-PMSI tunnel on PE-3 match the properties of the BIER tunnel, as follows:

```
*A:PE-3# show router 1 pim s-pmsi bier-root-addr 192.0.2.1
=====
PIM BIER Spmsi tunnels
=====
Root Addr          Sub-Domain ID BFR ID      MPLS Lbl
Multistream-ID     If-Index      Num Vpn SG's State
-----
192.0.2.1        0             1           524269
0                  73734       1           Up
=====
PIM BIER Spmsi interfaces : 1
=====
*A:PE-3#
```

PE-3 signals the join request for group 225.70.1.1 as an MP-BGP Source-Join update message to PE-1, so in turn PE-1 also creates a PIM state. On PE-1, the multicast traffic enters and leaves VPRN 1 via interfaces *int-PE-1-CE-1* and *mpls-if-73734* (S-PMSI), respectively, as follows:

```
*A:PE-1# show router 1 pim group 225.70.1.1 detail
=====
PIM Source Group ipv4
=====
Group Address      : 225.70.1.1
Source Address     : 10.1.1.11
RP Address         : 0
Advt Router        : 192.0.2.1
Flags              :                               Type           : (S,G)
Mode               : sparse
MRIB Next Hop     : 10.1.1.11
MRIB Src Flags     : direct
Keepalive Timer    : Not Running
Up Time           : 0d 00:29:23      Resolved By       : rtable-u

Up JP State        : Joined           Up JP Expiry      : 0d 00:00:00
Up JP Rpt          : Not Joined StarG  Up JP Rpt Override : 0d 00:00:00

Register State     : No Info
Reg From Anycast RP: No

Rpf Neighbor       : 10.1.1.11
Incoming Intf      : int-PE-1-CE-1
Outgoing Intf List : mpls-if-73731 (mpls-if-73734)

Curr Fwding Rate   : 65.912 kbps
Forwarded Packets  : 9594             Discarded Packets : 0
Forwarded Octets   : 14371812        RPF Mismatches    : 0
Spt threshold      : 0 kbps           ECMP opt threshold : 7
Admin bandwidth    : 1 kbps
-----
Groups : 1
=====
*A:PE-1#
```

The properties of the I-PMSI *mpls-if-73731* tunnel interface on PE-1 include the details for the joined multicast group, as follows:

```
*A:PE-1# show router 1 pim tunnel-interface "mpls-if-73731" detail
```

```

=====
PIM Interface ipv4 mpls-if-73731
=====
Admin Status      : Up                Oper Status      : Up
IPv4 Admin Status : Up                IPv4 Oper Status : Up
DR                : 192.0.2.1
Auto-created      : Yes
Transport Type    : MVPN-Pmsi

-----
PIM Group Source
-----
Group Address     : 225.70.1.1
Source Address    : 10.1.1.11
Interface         : mpls-if-73731   Type             : (S,G)
RP Address        : 0.0.0.0
Up Time          : 0d 00:33:59

Join Prune State  : Join              Expires           : Never
Prune Pend Expires : N/A

Assert State      : No Info

-----
Interfaces : 1
=====
*A:PE-1#

```

Because VRF 1 is configured to use selective provider tunnels, SR OS creates additional BIER tunnels when joining groups exceed the data threshold of 10 kb/s. With only group 225.70.1.1 joined, the BIER tunnels on PE-1 are as follows:

```

*A:PE-1# show router bier tunnel

=====
BIER Tunnels
=====
Tunnel-id      Type      Oper      No. Of Leaves
BFR Prefx     Bfr-ID   Mpls Label Sub-domain
-----
73731          tx       In service 2
192.0.2.1     1        524270    0

73732          rx       In service 0
192.0.2.3     3        524270    0

73733          rx       In service 0
192.0.2.4     4        524270    0

73734          tx       In service 1
192.0.2.1     1        524269    0

=====
BIER Tunnel entries : 4
=====
*A:PE-1#

```

The properties of the S-PMSI tunnel for group 225.70.1.1 on PE-1 show that the MPLS label 524269 and interface index 73734 are used, as follows. The label and interface index are the same as the highlighted BIER tunnel properties from the previous command.

```

*A:PE-1# show router 1 pim s-pmsi group-ip 225.70.1.1 detail

```



```

=====
PIM BIER Spmsi tunnels
=====
Root Address      : 192.0.2.1          BFR ID           : 1
Sub Domain       : 0                Mpls Lbl        : 524269
Multistream ID   : N/A              If Index        : 73734
Num VPN SGs     : 1                Oper State      : Up

VPN Group Address : 225.70.1.1
VPN Source Address : 10.1.1.11
State            : TX Joined        Mdt Threshold    : 10
Join Timer       : 0                Hold Down Timer  : 22
SG Age           : 2442             Expiry Timer     : 0
Threshold enabled : False          Receiver count   : 0
-----
PIM BIER Spmsi interfaces : 1
=====
*A:PE-1#

```

In summary, for the multicast traffic originated by CE-1 to VPRN 1, PE-1 passes the core network through a BIER tunnel and leaves VPRN 1, PE-3 to terminate in CE-3. A transit node in the core network (P-2 in the example of Figure 1) does not maintain any multicast state for the multicast stream.

With MVPN-enabled VPRN 1 instances defined on PE-1, PE-3, and PE-4, and group 225.70.1.1 active, an entry is added to the MVPN list. On PE-1, this list is as follows:

```

*A:PE-1# show router mvpn-list type bier

Legend: Sig = Signal  Pim-a = pim-asm  Pim-s = pim-ssm  A-D = Auto-Discovery
SR = Sender-Receiver  SO = Sender-Only  RO = Receiver-Only

=====
MVPN List
=====
VprnID   A-D   iPmsi/sPmsi  GroupAddr/Lsp-Template   IPv4(S,G)/(*,G)
          Sig   Mdt-Type
-----
1        Default  Bier/Bier    N/A                       1/0
          Bgp   SR
          0/0
-----
Total Mvpns : 1
=====

Total                PIM      RSVP      MLDP      BIER
-----
I-PMSI tunnels       0         0         0         1
TX S-PMSI tunnels    0         0         0         1
RX S-PMSI tunnels    0         0         0         0
RX PSEUDO S-PMSI tunnels 0         0         0         0
-----
Total IPv4 (S,G)/(*,G) : 1/0
Total IPv6 (S,G)/(*,G) : 0/0
=====
*A:PE-1#

```

As opposed to Rosen MVPNs, at no place in the network must multicast be configured in the base router. The status of PIM can be verified as follows, but for brevity the command is executed on P-2 only.

```

*A:P-2# show router pim status
MINOR: CLI PIM is not configured.

```

```
*A:P-2#
```

## Debug

The following debug configuration can be used for troubleshooting BIER:

```
debug
  router "Base"
    bier
      management
      template
      tunnel
    exit
  exit
exit
```

The log shows the trace when CE-4 joins group 225.70.1.1 on PE-4, VPRN 1. To see the interactions with IGMP, PIM, and BGP, these protocols should be debugged too, but these are omitted here. The message sequence indicates what happens when a new BIER tunnel is created on PE-4. The new BIER tunnel ID 73735 is highlighted.

```
*A:PE-4# show log log-id 1 ascending
```

```
=====  
Event Log 1  
=====
```

```
Description : (Not Specified)
```

```
Memory Log contents [size=100 next event=10 (not wrapped)]
```

```
1 2019/05/06 12:57:23.004 CEST MINOR: DEBUG #2001 Base BIER[TUNNEL inst 1]
"BIER[TUNNEL inst 1]: bierMttmProcessEvent
Process CREATE event for mttmIdx 73735"
```

```
2 2019/05/06 12:57:23.004 CEST MINOR: DEBUG #2001 Base BIER[TUNNEL inst 1]
"BIER[TUNNEL inst 1]: bierTunnelCreate
Create Tunnel 73735 , Type RX, PTA: isValid T, BFR ID 1, SD 0, PFX 192.0.2.1,
MPLS label 524269"
```

```
3 2019/05/06 12:57:23.004 CEST MINOR: DEBUG #2001 Base BIER[TUNNEL inst 1]
"BIER[TUNNEL inst 1]: bierTunnelPrefixAdd
Add Tunnel to PFX 192.0.2.4, SD 0 isLocal TRUE for tracking. Existing Tunnel
Tracked 1.Existing MVPN tracked 1. Existing Rx Tunnel Tracked 0"
```

```
4 2019/05/06 12:57:23.004 CEST MINOR: DEBUG #2001 Base BIER[TUNNEL inst 1]
"BIER[TUNNEL inst 1]: bierTunnelPrefixAdd
Add Tunnel to PFX 192.0.2.1, SD 0 isLocal FALSE for tracking. Existing Tunnel
Tracked 2.Existing MVPN tracked 2. Existing Rx Tunnel Tracked 0"
```

```
5 2019/05/06 12:57:23.004 CEST MINOR: DEBUG #2001 Base BIER[TUNNEL inst 1]
"BIER[TUNNEL inst 1]: bierTunnelTrackedSvcIdAdd
[Bier:1] ADD SvcId 2 Associated with Pfx 192.0.2.1. Existing TX/RX Tunnel 1/2"
```

```
6 2019/05/06 12:57:23.004 CEST MINOR: DEBUG #2001 Base BIER[TUNNEL inst 1]
"BIER[TUNNEL inst 1]: bierHandleTunnelCreate
Validate Tunnel 73735, subDomain 0, isPTAValid 1 ,BFR ID in PTA : 1, BFR ID
in DB : 1"
```

```
7 2019/05/06 12:57:23.004 CEST MINOR: DEBUG #2001 Base BIER[TUNNEL inst 1]
"BIER[TUNNEL inst 1]: bierTunnelUpdateFib
```

```
update FIB :- Tunnel 73735 Type : RX setId : 0 event TUN_ADD"

8 2019/05/06 12:57:23.004 CEST MINOR: DEBUG #2001 Base BIER[TUNNEL inst 1]
"BIER[TUNNEL inst 1]: bierTunnelProcessFibTunnelMsg
Process FIB msg TAPMAP_CHG for Tunnel 73735 OLD INFO : (IlmIdx: 0 p2mpIdx: 0,
mid :0, mcid :0, TapMap : 0x000000000000000000000000)NEW INFO : (IlmIdx: 24
p2mpIdx: 24, mid :42110, mcid :8381, TapMap : 0x000000000000000000000001)"

9 2019/05/06 12:57:23.004 CEST MINOR: DEBUG #2001 Base BIER[TUNNEL inst 1]
"BIER[TUNNEL inst 1]: bierTunnelMttmSendTunnelModify
Update MTTM for Tunnel 73735 with HW info and Oper state : UP"
*A:PE-4#
```

This new BIER tunnel is added to the list of tunnels, as follows:

```
*A:PE-4# show router bier tunnel

=====
BIER Tunnels
=====
Tunnel-id          Type          Oper          No. Of Leaves
BFR Prefix         Bfr-ID       Mpls Label    Sub-domain
-----
73731              tx            In service    2
192.0.2.4          4            524270        0

73732              rx            In service    0
192.0.2.3          3            524270        0

73733              rx            In service    0
192.0.2.1          1            524270        0

73735             rx          In service    0
192.0.2.1       1          524269      0
=====
BIER Tunnel entries : 4
=====
*A:PE-4#
```

## Conclusion

Using BIER as multicast transport technology provides operators an alternative to Rosen MVPN scenarios that optimizes network resources. Multicast provisioning is simplified because the use of PIM, mLDLP, or RSVP-TE as tunneling technologies can be avoided. Because no multicast control protocol is required in the core, no multicast state must be maintained; consequently, there is no need to reconverge state and resignal multicast state in the case of a network failure.

---

## Layer 3 VPN: VPRN Type Spoke

This chapter provides information about Layer 3 VPRN CE hub and spoke architecture.

Topics in this chapter include:

- [Applicability](#)
- [Overview](#)
- [Configuration](#)
- [Conclusion](#)

### Applicability

This chapter was initially written for SR OS Release 12.0. However, the CLI in the current edition is based on SR OS Release 22.2.R1.

Knowledge of Nokia's Layer 3 VPN concepts is assumed throughout this document.

### Overview

This chapter provides a basic technology overview and configuration examples of a network topology used for a Layer 3 VPRN CE hub and spoke architecture.

### VPRN type hub

In SR OS Releases earlier than 12.0, a CE hub and spoke architecture was partially supported. Internal optimization was available for the hub sites connected to the same PE router only. This feature is known as VPRN type hub. If, on the other hand, multiple spoke sites were connected to the same PE router, separate VPRN instances had to be created to maintain the split horizon forwarding behavior. This approach was complex, hard to maintain and consumed extra VPRN instances.

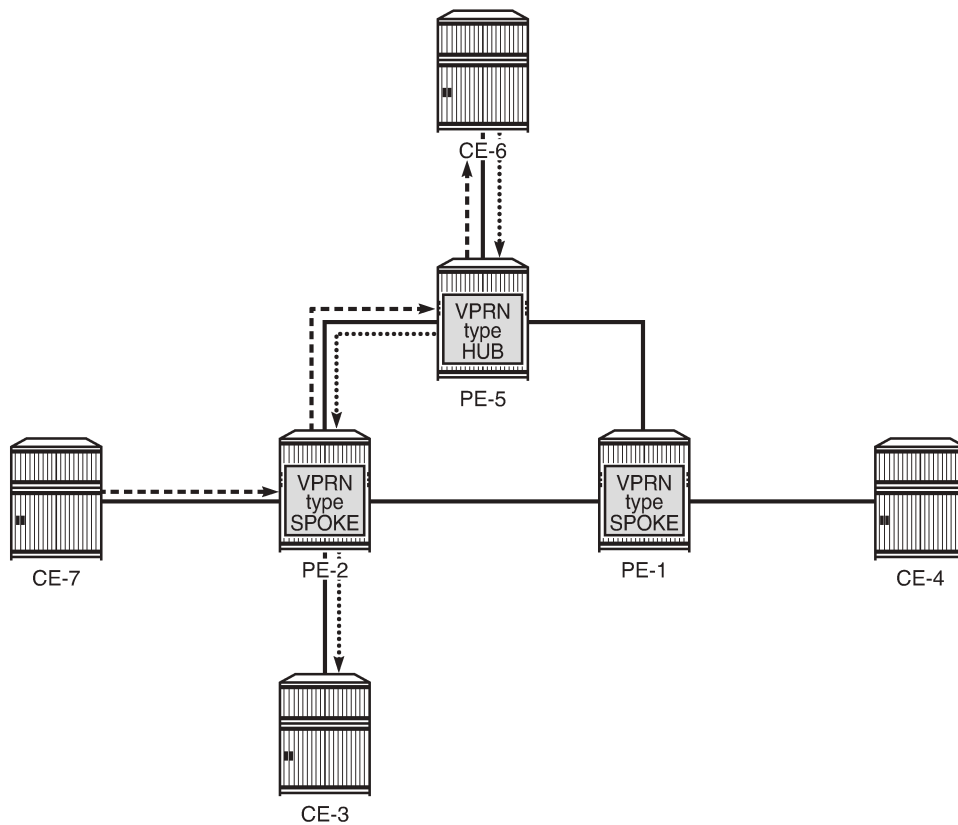
### VPRN type spoke

Release 12.0.R1 added new functionality to overcome these limitations. Introducing the VPRN type spoke feature allows multiple spoke sites to be kept within the same VPRN instance while at the same time maintaining the split horizon approach such that spoke sites cannot send traffic directly to each other.

The primary goal of the feature is to allow multiple spoke sites to be part of a single VPRN instance without allowing direct communication between the spoke CE sites which are part of that VPRN (of type spoke).

The packet flow is demonstrated in [Figure 25: CE hub and spoke data path](#).

Figure 25: CE hub and spoke data path



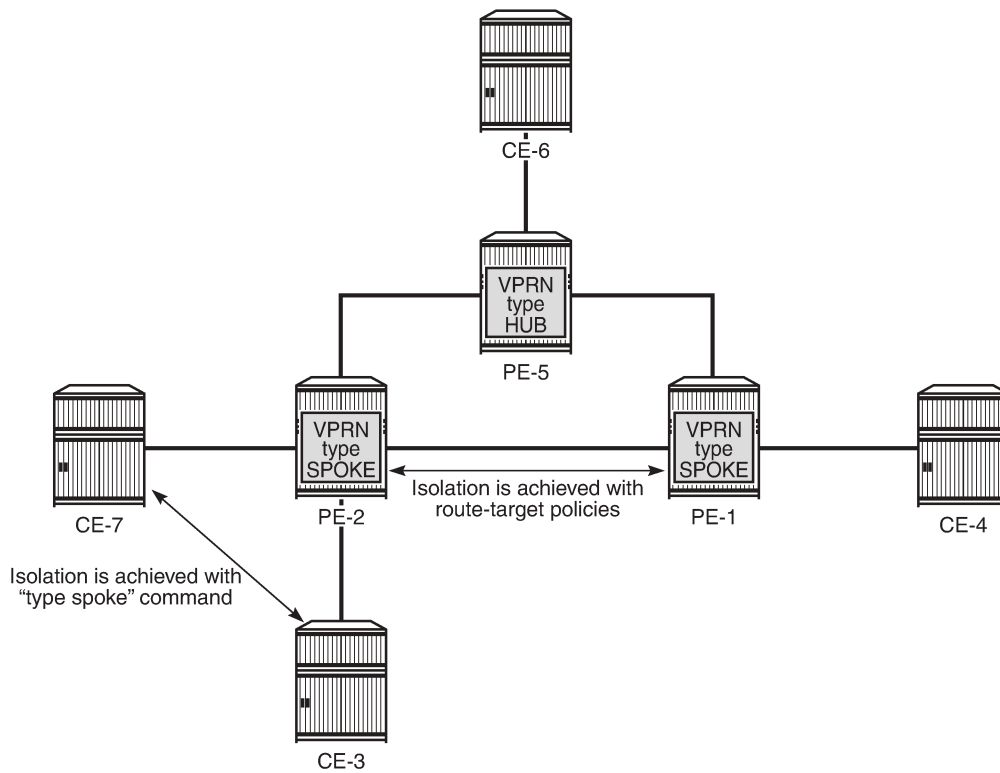
25460

The only way for CE-7 to communicate with CE-3 is via hub site CE-6. The same applies to the communication between CE-7 and CE-4. The VPRN on PE-2 is configured as **type spoke** and has IP interfaces using SAPs or spoke SDPs that are considered spoke sites only. No direct communication between any of the spoke CE sites in the network is allowed.

Direct communication between the spoke CE sites is blocked using two techniques, as illustrated in [Figure 26: CE hub and spoke control plane isolation](#).

- Using the **type spoke** command under the **vprn** context as explained later.
- The extended community configuration using route-target policies (this is not covered in detail in this chapter).

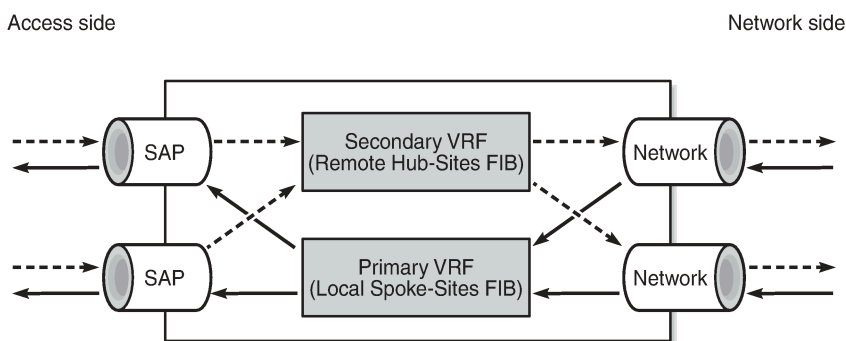
Figure 26: CE hub and spoke control plane isolation



25461

When a VPRN on a PE router is configured as **type spoke**, then the internal forwarding logic changes as demonstrated in [Figure 27: Internal VPRN logic on a PE router](#).

Figure 27: Internal VPRN logic on a PE router



25462

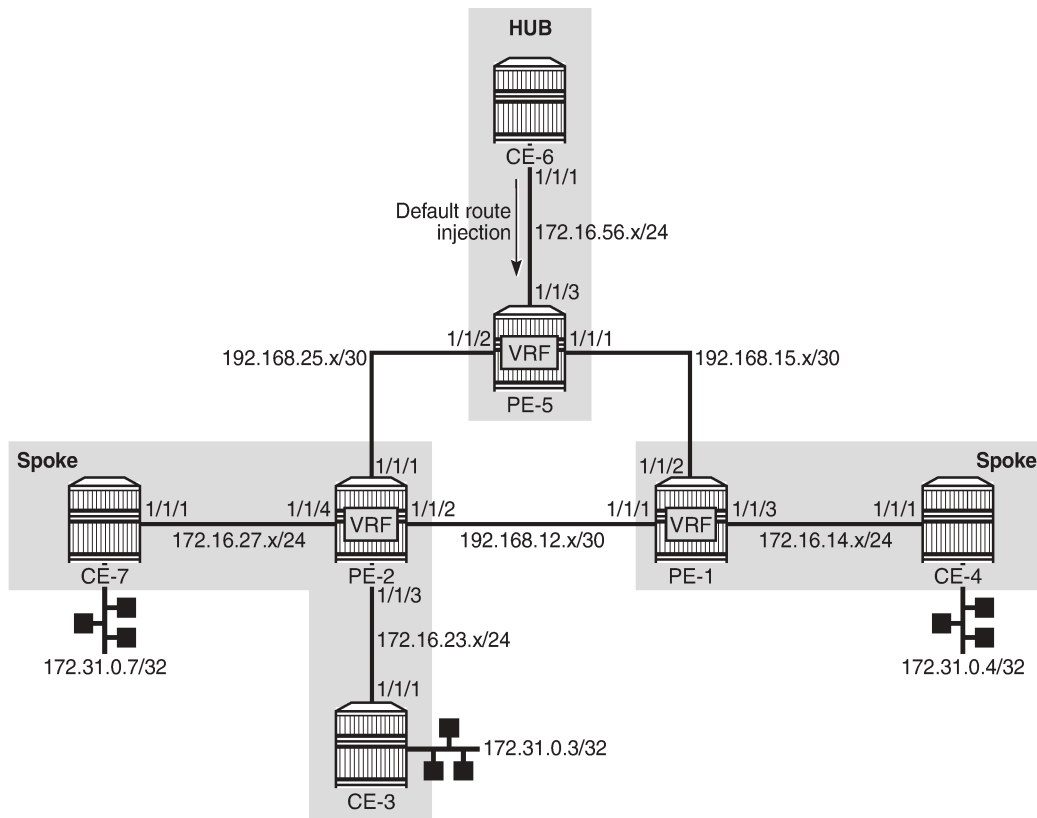
- VPRNs of type spoke create a primary and a secondary VRF internally to the VPRN:
  - The primary VRF is used for forwarding traffic from the network interfaces toward the IP interfaces using SAPs or spoke SDPs. This VRF is populated with routes learned from the spoke CE sites connected to the local PE through IP interfaces using SAPs or spoke SDPs.

- The secondary VRF is used for forwarding traffic from the IP interfaces using SAPs or spoke SDPs toward the network interfaces or other VPRN instances. This VRF is populated with routes learned via MP-BGP from hub sites.
- VPRNs of type spoke export routes using a specific extended community (for instance, spoke-ext-comm) via an export policy to identify them as spoke site originated routes.
  - This community is not hard-coded and has to be configured manually (see configuration example later).
- VPRNs of type spoke import routes (using an import policy) received from other PEs or VPRN instances with a hub specific community only (for example, hub-ext-comm). Routes with spoke-ext-comm community are ignored.
  - This community is not hard-coded and has to be configured manually (see configuration example later).
- Multiple VPRNs of type spoke and hub can coexist on the same PE if they use different VPRN instances.
- The configuration of type hub and type spoke is mutually exclusive within one VPRN instance.

## Configuration

The physical topology and addressing scheme are presented in [Figure 28: CE hub and spoke topology and addressing scheme](#).

Figure 28: CE hub and spoke topology and addressing scheme



25463

The configuration of PE-2 and PE-5 are the main focus of this example. The configuration of PE-1 is similar to that of PE-2.

## Hub site configuration

Only the essential part of the configuration is provided for the hub site.

Vrf-import and vrf-export policies are used to manipulate the vrf-target in order to achieve logical isolation between the spoke sites in the network.

```
# on PE-5:
configure
router Base
  policy-options
  begin
  community "hub-ext-comm"
    members "target:64500:11"
  exit
  community "spoke-ext-comm"
    members "target:64500:12"
  exit
  policy-statement "vrf-export"
    default-action accept
```



```

        community add "hub-ext-comm"
    exit
exit
policy-statement "vrf-import"
    entry 10
        from
            community "spoke-ext-comm"
        exit
        action accept
    exit
    default-action drop
exit
policy-statement "export-ospf"
    entry 10
        from
            protocol direct
        exit
        action accept
    exit
    default-action accept
exit
commit
exit

```

PE-5 is configured with VPRN 1 providing OSPF connectivity to customer CE-6.

```

# on PE-5:
configure
    service
        vprn 1 name "VPRN1" customer 1 create
        description "VPRN type hub"
        type hub
        interface "int-PE-5-CE-6" create
            address 172.16.56.1/24
            sap 1/1/3:1 create
        exit
    exit
    bgp-ipvpn
        mpls
            auto-bind-tunnel
                resolution any
            exit
            route-distinguisher 64500:15
            vrf-import "vrf-import"
            vrf-export "vrf-export"
            no shutdown
        exit
    exit
    ospf
        export "export-ospf"
        area 0.0.0.0
            interface "int-PE-5-CE-6"
                interface-type point-to-point
                mtu 1500
                no shutdown
            exit
        exit
    exit
    no shutdown
exit

```

```
no shutdown
exit
```

At the same time, CE-6 is configured to advertise a default route which is used by all remote spoke CE sites to forward traffic via CE-6.

```
# on CE-6:
configure
  router Base
    policy-options
      begin
        policy-statement "export-ospf-default"
          entry 10
            from
              protocol static
            exit
            action accept
            exit
          exit
        exit
      exit
    commit
  exit
```

```
# on CE-6:
configure
  service
    vprn 1 name "VPRN1" customer 1 create
      interface "int-CE-6-PE-5" create
        address 172.16.56.2/24
        sap 1/1/1:1 create
        exit
      exit
    static-route-entry 0.0.0.0/0
      black-hole
      no shutdown
      exit
    exit
    ospf 192.0.2.6
      export "export-ospf-default"
      ignore-dn-bit
      suppress-dn-bit
      area 0.0.0.0
        interface "int-CE-6-PE-5"
          interface-type point-to-point
          mtu 1500
          no shutdown
        exit
      exit
    exit
  no shutdown
  exit
  no shutdown
exit
```

## Spoke site configuration

According to the example topology, two spoke VPRNs are present: one VPRN with two CE spoke sites connected is located on PE-2, and another VPRN with one spoke CE site on PE-1. The service configuration for PE-2 is as follows with the one for PE-1 being similar.

Vrf-import and vrf-export policies are used to build a hub-and-spoke topology in order to achieve a logical isolation between spoke sites connected to different PE routers.

```
# on PE-2:
configure
  router Base
    policy-options
      begin
        community "hub-ext-comm"
          members "target:64500:11"
        exit
        community "spoke-ext-comm"
          members "target:64500:12"
        exit
        policy-statement "vrf-export"
          default-action accept
          community add "spoke-ext-comm"
        exit
      exit
    policy-statement "vrf-import"
      entry 10
        from
          community "hub-ext-comm"
        exit
        action accept
      exit
    exit
    default-action drop
  exit
  policy-statement "export-ospf"
    default-action accept
  exit
  exit
  commit
exit
```

PE-2 is configured with VPRN 1, which has OSPF connectivity to the customer CE-3 and CE-7. The **type spoke** command is used to prevent direct CE spoke to CE spoke communications for this VPRN.

```
# on PE-2:
configure
  service
    vprn 1 name "VPRN1" customer 1 create
    description "VPRN type spoke"
    type spoke
    interface "int-PE-2-CE-3" create
      address 172.16.23.1/24
      sap 1/1/3:1 create
    exit
  exit
  interface "int-PE-2-CE-7" create
    address 172.16.27.1/24
    sap 1/1/4:1 create
  exit
  exit
  bgp-ipvpn
    mpls
      auto-bind-tunnel
        resolution any
    exit
    route-distinguisher 64500:12
```

```

        vrf-import "vrf-import"
        vrf-export "vrf-export"
        no shutdown
    exit
exit
ospf
    export "export-ospf"
    area 0.0.0.0
        interface "int-PE-2-CE-3"
            interface-type point-to-point
            mtu 1500
            no shutdown
        exit
        interface "int-PE-2-CE-7"
            interface-type point-to-point
            mtu 1500
            no shutdown
        exit
    exit
    no shutdown
exit
no shutdown

```

For connectivity verification purposes, CE-3, CE-4, and CE-7 are configured to advertise their internal loopback interfaces via OSPF:

- CE-3 advertises 172.31.0.3/32
- CE-4 advertises 172.31.0.4/32
- CE-7 advertises 172.31.0.7/32

```

# on CE-3:
configure
    service
        vprn 1 name "VPRN1" customer 1 create
        interface "int-CE-3-PE-2" create
            address 172.16.23.2/24
            sap 1/1/1:1 create
            exit
        exit
        interface "lo0" create
            address 172.31.0.3/32
            loopback
        exit
        ospf 192.0.2.3
            ignore-dn-bit
            suppress-dn-bit
            area 0.0.0.0
                interface "int-CE-3-PE-2"
                    interface-type point-to-point
                    mtu 1500
                exit
                interface "lo0"
            exit
        exit
        no shutdown
    exit
    no shutdown
exit

```

## Hub site verification

The Routing Information Base (RIB) for VPRN 1 on hub site PE-5 lists all reachable networks:

```
*A:PE-5# show router 1 route-table

=====
Route Table (Service: 1)
=====
Dest Prefix[Flags]                               Type   Proto   Age           Pref
  Next Hop[Interface Name]                       Metric
-----
0.0.0.0/0                                         Remote OSPF     00h02m32s  150
  172.16.56.2                                     1
172.16.14.0/24                                    Remote BGP VPN 00h01m07s  170
  192.0.2.1 (tunneled)                            10
172.16.14.1/32                                    Remote BGP VPN 00h01m07s  170
  192.0.2.1 (tunneled)                            10
172.16.23.0/24                                    Remote BGP VPN 00h01m01s  170
  192.0.2.2 (tunneled)                            10
172.16.23.1/32                                    Remote BGP VPN 00h01m01s  170
  192.0.2.2 (tunneled)                            10
172.16.27.0/24                                    Remote BGP VPN 00h01m01s  170
  192.0.2.2 (tunneled)                            10
172.16.27.1/32                                    Remote BGP VPN 00h01m01s  170
  192.0.2.2 (tunneled)                            10
172.16.56.0/24                                    Local  Local   00h03m57s   0
  int-PE-5-CE-6                                   0
172.31.0.3/32                                     Remote BGP VPN 00h01m01s  170
  192.0.2.2 (tunneled)                            10
172.31.0.4/32                                     Remote BGP VPN 00h00m40s  170
  192.0.2.1 (tunneled)                            10
172.31.0.7/32                                     Remote BGP VPN 00h00m30s  170
  192.0.2.2 (tunneled)                            10
-----
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 forwarding table (FIB) for the primary VRF of VPRN 1 is displayed using following command. All remote spoke and hub sites are reachable via this VRF.

```
*A:PE-5# show router 1 fib 1

=====
FIB Display
=====
Prefix [Flags]                               Protocol
  NextHop
-----
0.0.0.0/0                                         OSPF
  172.16.56.2 (int-PE-5-CE-6)
172.16.14.0/24                                    BGP_VPN
  192.0.2.1 (VPRN Label:524284 Transport:LDP)
172.16.14.1/32                                    BGP_VPN
  192.0.2.1 (VPRN Label:524284 Transport:LDP)
172.16.23.0/24                                    BGP_VPN
  192.0.2.2 (VPRN Label:524284 Transport:LDP)
```

```

172.16.23.1/32                                BGP_VPN
  192.0.2.2 (VPRN Label:524284 Transport:LDP)
172.16.27.0/24                                BGP_VPN
  192.0.2.2 (VPRN Label:524284 Transport:LDP)
172.16.27.1/32                                BGP_VPN
  192.0.2.2 (VPRN Label:524284 Transport:LDP)
172.16.56.0/24                                LOCAL
  172.16.56.0 (int-PE-5-CE-6)
172.31.0.3/32                                 BGP_VPN
  192.0.2.2 (VPRN Label:524284 Transport:LDP)
172.31.0.4/32                                 BGP_VPN
  192.0.2.1 (VPRN Label:524284 Transport:LDP)
172.31.0.7/32                                 BGP_VPN
  192.0.2.2 (VPRN Label:524284 Transport:LDP)
-----
Total Entries : 11
-----
=====

```

The forwarding table for the secondary VRF of VPRN 1 is displayed using following command, including the **secondary** keyword. All local hub CE sites are reachable via this VRF.

```

*A:PE-5# show router 1 fib 1 secondary
=====
FIB Display
=====
Prefix [Flags]                                Protocol
NextHop
-----
0.0.0.0/0                                     OSPF
  172.16.56.2 (int-PE-5-CE-6)
172.16.56.0/24                                LOCAL
  172.16.56.0 (int-PE-5-CE-6)
-----
Total Entries : 2
-----
=====

```

## Spoke site verification

The RIB for VPRN 1 on PE-2 (spoke VPRN) lists all reachable networks.

The other spoke sites connected to the remote PEs are not present in the routing table, in this example, CE-4 with prefixes such as 172.31.0.4/32 and 172.16.14.0/24.

The local interface addresses of PE-2 (172.16.23.1/32 and 172.16.27.1/32) are present in the routing table of VPRN 1, as follows. From a FIB point of view, these are reachable from any spoke VPRN, but the spoke CE's router host addresses are not. This fact does not influence the data plane isolation for the customer networks.

```

*A:PE-2# show router 1 route-table
=====
Route Table (Service: 1)
=====
Dest Prefix[Flags]                            Type   Proto   Age           Pref
Next Hop[Interface Name]                    Metric
-----
0.0.0.0/0                                     Remote BGP VPN 00h22m49s 170

```

```

192.0.2.5 (tunneled)
172.16.23.0/24          Local  Local  00h22m53s  0
  int-PE-2-CE-3
172.16.23.1/32         Local  Host   00h22m53s  0
  int-PE-2-CE-3
172.16.27.0/24         Local  Local  00h22m53s  0
  int-PE-2-CE-7
172.16.27.1/32         Local  Host   00h22m53s  0
  int-PE-2-CE-7
172.16.56.0/24         Remote BGP VPN 00h22m49s 170
  192.0.2.5 (tunneled)
172.16.56.1/32         Remote BGP VPN 00h22m49s 170
  192.0.2.5 (tunneled)
172.31.0.3/32          Remote OSPF  00h22m40s 10
  172.16.23.2
172.31.0.7/32          Remote OSPF  00h22m24s 10
  172.16.27.2
-----
No. of Routes: 9
Flags: n = Number of times nexthop is repeated
      B = BGP backup route available
      L = LFA nexthop available
      S = Sticky ECMP requested
=====

```

The FIB for the primary VRF of VPRN 1 shows all local spoke sites are reachable via this VRF, as follows:

```

*A:PE-2# show router 1 fib 1
=====
FIB Display
=====
Prefix [Flags]          Protocol
NextHop
-----
172.16.23.0/24          LOCAL
  172.16.23.0 (int-PE-2-CE-3)
172.16.23.1/32         HOST
  Blackhole
172.16.27.0/24         LOCAL
  172.16.27.0 (int-PE-2-CE-7)
172.16.27.1/32         HOST
  Blackhole
172.31.0.3/32          OSPF
  172.16.23.2 (int-PE-2-CE-3)
172.31.0.7/32          OSPF
  172.16.27.2 (int-PE-2-CE-7)
-----
Total Entries : 6
=====

```

The FIB for the secondary VRF of VPRN 1 shows the remote hub site (address 172.16.56.0/24) is reachable via this VRF, as follows:

```

*A:PE-2# show router 1 fib 1 secondary
=====
FIB Display
=====
Prefix [Flags]          Protocol
NextHop
-----

```

```

-----
0.0.0.0/0                                     BGP_VPN
  192.0.2.5 (VPRN Label:524284 Transport:LDP)
172.16.23.1/32                               HOST
  Blackhole
172.16.27.1/32                               HOST
  Blackhole
172.16.56.0/24                               BGP_VPN
  192.0.2.5 (VPRN Label:524284 Transport:LDP)
172.16.56.1/32                               BGP_VPN
  192.0.2.5 (VPRN Label:524284 Transport:LDP)
-----
Total Entries : 5
=====

```

## Spoke sites connectivity verification

Without the VPRN spoke type configuration in VPRN 1 on PE-2, CE-3 takes the shortest path to CE-7, which violates the hub-and-spoke design approach explained earlier.

The VPRN has to be disabled (**shutdown**) in order to modify the type.

```
*A:PE-2>config>service>vprn# no type
INFO: PIP #1162 Instance must be 'shutdown'
```

```
# on PE-2:
configure
service
  vprn "VPRN1"
  shutdown
  no type
  no shutdown
```



### Note:

In this setup, a VPRN is configured on CE-3, but that is not necessary.

Traffic from CE-3 takes the shortest path to CE-7, because VPRN 1 on PE-2 is not configured with spoke type anymore.

```
*A:CE-3# traceroute router 1 172.31.0.7 no-dns
traceroute to 172.31.0.7, 30 hops max, 40 byte packets
 1 172.16.23.1  2.74 ms  2.60 ms  2.64 ms
 2 172.31.0.7   3.28 ms  3.21 ms  3.07 ms
```

After enabling the **type spoke** feature on PE-2, CE-3 takes the longest path via hub CE-6 to reach CE-7, as it should.

```
# on PE-2:
configure
service
  vprn "VPRN1"
  shutdown
  type spoke
```

```
*A:CE-3# traceroute router 1 172.31.0.7 no-dns
```



```
traceroute to 172.31.0.7, 30 hops max, 40 byte packets
 1 172.16.23.1    2.55 ms  2.80 ms  2.66 ms
 2 0.0.0.0      * * *
 3 172.16.56.2    4.47 ms  4.36 ms  4.33 ms
 4 172.16.56.1    4.49 ms  4.17 ms  4.20 ms
 5 172.16.27.1    4.10 ms  4.22 ms  3.98 ms
 6 172.31.0.7     6.63 ms  6.42 ms  6.61 ms
```

Similarly, the long path is taken by CE-3 to reach CE-4, as follows. This is unrelated to the VPRN type. It is achieved by policies.

```
*A:CE-3# traceroute router 1 172.31.0.4 no-dns
traceroute to 172.31.0.4, 30 hops max, 40 byte packets
 1 172.16.23.1    2.31 ms  2.47 ms  2.55 ms
 2 0.0.0.0      * * *
 3 172.16.56.2    4.53 ms  4.57 ms  4.35 ms
 4 172.16.56.1    4.66 ms  4.46 ms  4.47 ms
 5 172.16.14.1    6.06 ms  6.21 ms  6.08 ms
 6 172.31.0.4     7.00 ms  6.87 ms  7.30 ms
```

## Conclusion

The VPRN type spoke feature completes the CE hub and spoke solution. It brings an additional level of simplicity, scalability, and flexibility to operators using this VPRN architecture for their customers.

# NG-MVPN Configuration with MPLS

This chapter provides information about NG-MVPN configuration with MPLS.

Topics in this chapter include:

- [Applicability](#)
- [Overview](#)
- [Configuration](#)
- [Conclusion](#)

## Applicability

This chapter was initially written for SR OS Release 9.0.R5. The CLI in this edition corresponds to Release 23.7.R2. There are no prerequisites for this configuration.

## Overview

Multicast VPN (MVPN) or Next Generation IP Multicast in an IP-VPN (NG-MVPNs) architectures describe a set of virtual routing and forwarding (VRFs) or virtual private routed networks (VPRNs) that support the transport of multicast traffic across a provider network. MVPNs are defined in RFC 6513, *Multicast in MPLS/BGP IP VPNs*, and RFC 6514, *BGP Encodings and Procedures for Multicast in MPLS/IP VPNs*.

Initial MVPN deployments were originally based on Rosen MVPN (RFC 6037) which described the protocols and procedures required to support an IP Multicast VPN. There were a number of limitations with the Rosen MVPN implementation including, but not limited to:

- Rosen MVPN requires a set of multicast distribution trees (MDTs) per VPN, which requires a PIM state per MDT. There is no option to aggregate MDTs across multiple VPNs.
- Initially, customer signaling, PE discovery and Data MDT signaling were all PIM-based because there was no mechanism available to decouple these. Now, PE discovery is supported using a BGP MDT address family identifier/subsequent address family identifier (AFI/SAFI), however, the data MDT still needs PIM.
- There is no mechanism for using MPLS to encapsulate multicast traffic in the VPN. GRE is the only encapsulation method available in Rosen MVPN.
- Rosen MVPN multicast trees are signaled using PIM only. MVPN allows the use of mLDP and RSVP P2MP LSPs.
- PE to PE protocol exchanges for Rosen MVPN is achieved using PIM only. MVPN allows for the use of BGP signaling as per unicast Layer 3 VPNs.

NG-MVPN addresses these limitations by extending the idea of the per-VRF tree by introducing the idea of provider multicast service interfaces (PMSIs). These are equivalent to the default MDTs of Rosen MVPN. NG-MVPN allows the decoupling of the mechanisms required to create a multicast VPN, such as PE auto-discovery (which PEs are members of which VPN), PMSI signaling (creation of tunnels between PEs), and

customer multicast signaling (multicast signaling -IGMP/PIM- received from customer edge routers). Two types of PMSI exist:

- Inclusive (I-PMSI) contains all the PEs for an MVPN; I-PMSI is the default multicast data path between all PEs of the same VPN.
- Selective (S-PMSI): contains only a subset of PEs of an MVPN; used to optimize multicast stream distribution to only the PEs with active receivers for those streams.

The [NG-MVPN Configuration with PIM](#) chapter contains the VPN configuration required for the provider multicast domain using PIM Any Source Multicast (ASM) with auto-discovery based on PIM or BGP auto-discovery (AD), PIM used for the customer multicast signaling and PIM Source Specific Multicast (SSM) used for the S-PMSI creation. The customer domain configuration covers the following cases:

- PIM ASM with the Rendezvous Point (RP) in the provider PE
- PIM ASM using anycast RP on the provider RPs
- PIM SSM

This chapter introduces some of the features that were not supported at the time of writing of chapter [NG-MVPN Configuration with PIM](#) (Release 7.0). It provides configuration details to implement:

- Multicast LDP (mLDP) and RSVP-TE Point to Multipoint (P2MP) for building customer trees (C-trees) which are using MPLS instead of PIM techniques.
- MVPN source redundancy.
- MDT AFI/SAFI (to fully interoperate with Cisco networks).

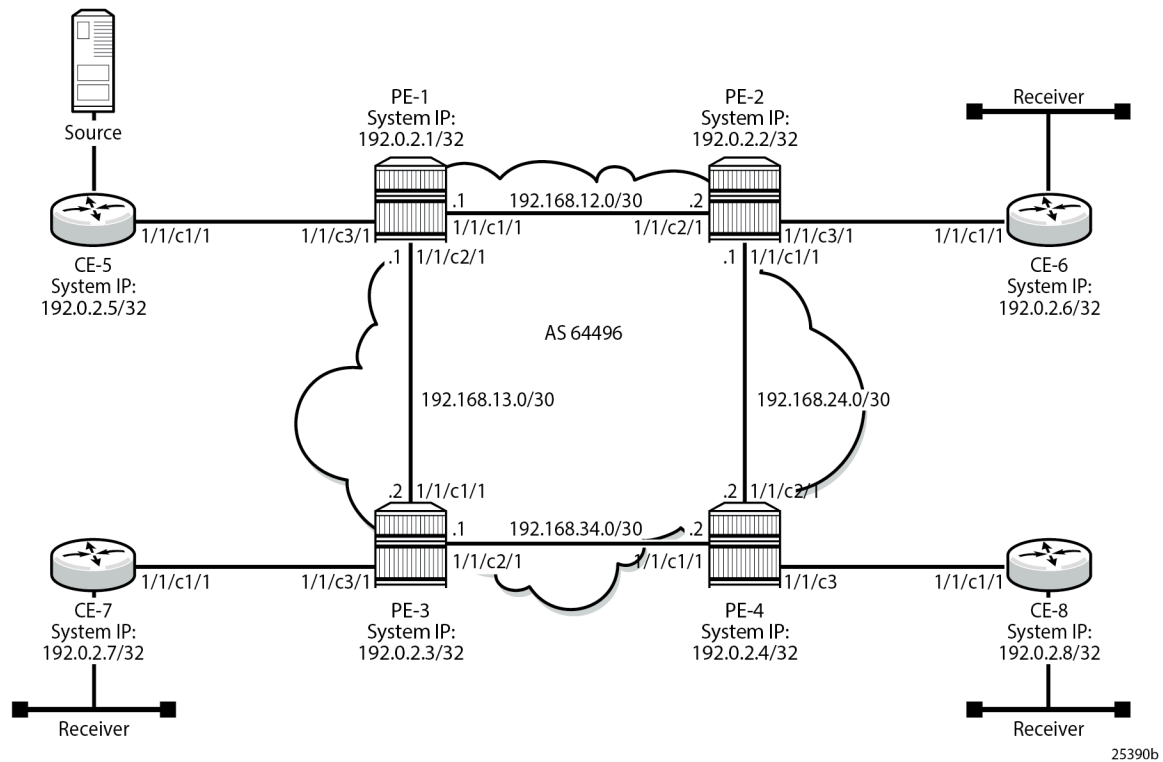
PIM SSM is the only case addressed in this example, other PIM customer domain configurations are out of the scope. For more information see: [NG-MVPN Configuration with PIM](#).

The network topology is shown in [Figure 29: Network Topology](#). The setup consists of four 7750 SRs acting as provider edge (PE) routers within a single autonomous system (AS).

- Full mesh IS-IS in the AS (OSPF could be used instead)
- LDP on all interfaces in each AS (RSVP could be used instead)
- MP-iBGP sessions between the PE routers in the AS (route reflectors (RRs) could also be used).
- Layer 3 VPN on all PEs with identical route targets.

Connected to each PE is a single 7750 SR acting as a customer edge (CE) router. CE-5 has a multicast source connected, and CE-6, CE-7, and CE-8 each have a single receiver connected which receive the multicast streams from the source. In this setup, each receiver is IGMPv3 capable. If the receiver is IGMPv3 capable, it issues IGMPv3 reports that may include a list of required source addresses.

Figure 29: Network Topology



When the receiver wants to become a member of any group, the source address of the group must be known to the CE. As a result, the source address must be IP reachable by each CE, so it is advertised using BGP by CE-5 to the PEs with attachment circuits in the VPRN. Static routes are then configured on the receiver CEs to achieve IP reachability to the source address of the multicast group.

Multicast traffic from the source is streamed toward router CE-5. Receivers connected to CE-6, CE-7, and CE-8 are interested in joining this multicast group.

The CEs are PIM enabled routers, which form a PIM adjacency with their nearest PE. Between the PEs across the provider network, there are no PIM adjacencies, because BGP auto-discovery and BGP signaling are used. Selective PMSI using mLDP or RSVP P2MP are out of the scope of this chapter. Selective PMSI using PIM SSM is supported too. I-PMSI and S-PMSI must use the same tunneling technology, either PIM/GRE, or mLDP, or RSVP P2MP.

## Configuration

The configuration is divided into the following sections:

- [Provider Common Configuration](#)
  - [PE Global Configuration](#)
- [PE VPRN Configuration and PE VPRN Multicast Configuration for NG-MVPN](#)
  - [PMSI using mLDP](#)

- PMSI using RSVP-TE
- MVPN Source Redundancy
- PE VPRN Configuration and PE VPRN Multicast Configuration for Rosen MVPN using MDT AFI SAFI

## Provider Common Configuration

### PE Global Configuration

This section describes the common configuration required for each PE within the provider multicast domain, regardless of the MVPN PE auto-discovery or customer signaling methods. This includes interior gateway protocol (IGP) and VPRN service configuration.

The configuration tasks can be summarized as follows:

- PE global configuration.  
This includes configuration of the IGP (IS-IS is used); configuration of link layer LDP between PEs (LDP is used here as the method to interconnect VPRNs); configuration of iBGP between PEs to facilitate VPRN route learning.
- VPRN configuration on the PEs.  
This includes configuration of basic VPRN parameters (route-distinguisher, route target communities), configuration of attachment circuits toward CEs, configuration of VRF routing protocol and any routing policies.
- PIM within the VRF and MVPN parameters - I-PMSI
- CE configuration.

#### Step 1.

Configure the interfaces, the IGP (IS-IS) in all PE nodes (where IS-IS redistributes route reachability to all routers) and LDP in the interfaces (link layer LDP). To facilitate the IS-IS configuration, all routers are Level2-Level1 capable within the same ISIS area-id, so there is only a single topology area in the network (all routers share the same topology). The configuration for PE-1 is displayed below.

```
# on PE-1
configure
router
  interface "int-PE-1-PE-2"
    address 192.168.12.1/30
    port 1/1/c1/1
    no shutdown
  exit
  interface "int-PE-1-PE-3"
    address 192.168.13.1/30
    port 1/1/c2/1
    no shutdown
  exit
  interface "system"
    address 192.0.2.1/32
    no shutdown
  exit
  autonomous-system 64496
  isis 0
    area-id 49.0001
```

```

traffic-engineering
interface "system"
    passive
    no shutdown
exit
interface "int-PE-1-PE-2"
    interface-type point-to-point
    no shutdown
exit
interface "int-PE-1-PE-3"
    interface-type point-to-point
    no shutdown
exit
no shutdown
exit
ldp
interface-parameters
    interface "int-PE-1-PE-2" dual-stack
        ipv4
        no shutdown
    exit
    no shutdown
    exit
    interface "int-PE-1-PE-3" dual-stack
        ipv4
        no shutdown
    exit
    no shutdown
    exit
exit
no shutdown
exit
exit all
    
```

The configuration for the rest of nodes is similar. The IP addresses can be derived from [Figure 29: Network Topology](#).

### Step 2.

Verify that IS-IS adjacencies and LDP peer sessions are formed.

```
*A:PE-1# show router isis adjacency
```

```

=====
Rtr Base ISIS Instance 0 Adjacency
=====
System ID                Usage State Hold Interface          MT-ID
-----
PE-2                      L1L2 Up    27   int-PE-1-PE-2          0
PE-3                      L1L2 Up    27   int-PE-1-PE-3          0
-----
Adjacencies : 2
=====
    
```

```
*A:PE-1# show router ldp session ipv4
```

```

=====
LDP IPv4 Sessions
=====
Peer LDP Id              Adj Type  State           Msg Sent  Msg Recv  Up Time
-----
192.0.2.2:0              Link      Established     39        40        0d 00:01:21
192.0.2.3:0              Link      Established     30        31        0d 00:00:55
    
```

```
-----
No. of IPv4 Sessions: 2
=====
```

### Step 3.

Configure iBGP full mesh between the PEs for VPRN routing (Route Reflectors could also be an option).

```
# on PE-1
configure
router
  bgp
    rapid-withdrawal
    rapid-update mvpn-ipv4 mdt-safi
    group "INTERNAL"
      family vpn-ipv4 mvpn-ipv4 mdt-safi
      type internal
      neighbor 192.0.2.2
      exit
      neighbor 192.0.2.3
      exit
      neighbor 192.0.2.4
      exit
    exit
  no shutdown
exit all
```

The families configured under the group "INTERNAL" are vpn-ipv4, mvpn-ipv4, and mdt-safi, because the three families are referenced in this chapter.

The mdt-safi parameter is not needed for NG-MVPN (mLDP/RSVP scenarios) and is only required for Rosen MVPN with MDT AFI SAFI.

Rapid withdrawal (configured on all PEs) disables the minimum route advertisement interval (MRAI) on sending BGP withdrawals. Rapid update (configured for MVPN-IPv4 and MDT AFI/SAFI address families) disables the MRAI interval on sending BGP update messages for the address family MVPN-IPv4 and MDT-SAFI).

### Step 4.

Verify that BGP peer relationships are established.

```
*A:PE-1# show router bgp summary
=====
BGP Router ID:192.0.2.1      AS:64496      Local AS:64496
=====
BGP Admin State      : Up          BGP Oper State      : Up
Total Peer Groups    : 1           Total Peers         : 3
Total VPN Peer Groups : 0           Total VPN Peers     : 0
Current Internal Groups : 1       Max Internal Groups : 1
Total BGP Paths      : 21          Total Path Memory   : 7392

Total IPv4 Remote Rts : 0           Total IPv4 Rem. Active Rts : 0
Total IPv6 Remote Rts : 0           Total IPv6 Rem. Active Rts : 0
---snip---

=====
BGP Summary
=====
Legend : D - Dynamic Neighbor
=====
Neighbor
Description
```

	AS	PktRcvd	InQ	Up/Down	State	Rcv/Act/Sent (Addr Family)
		PktSent	OutQ			
192.0.2.2	64496	5	0	00h00m31s	0/0/0 (VpnIPv4)	
		5	0		0/0/0 (MvpnIPv4)	
					0/0/0 (MdtSafi)	
192.0.2.3	64496	4	0	00h00m23s	0/0/0 (VpnIPv4)	
		4	0		0/0/0 (MvpnIPv4)	
					0/0/0 (MdtSafi)	
192.0.2.4	64496	4	0	00h00m15s	0/0/0 (VpnIPv4)	
		4	0		0/0/0 (MvpnIPv4)	
					0/0/0 (MdtSafi)	

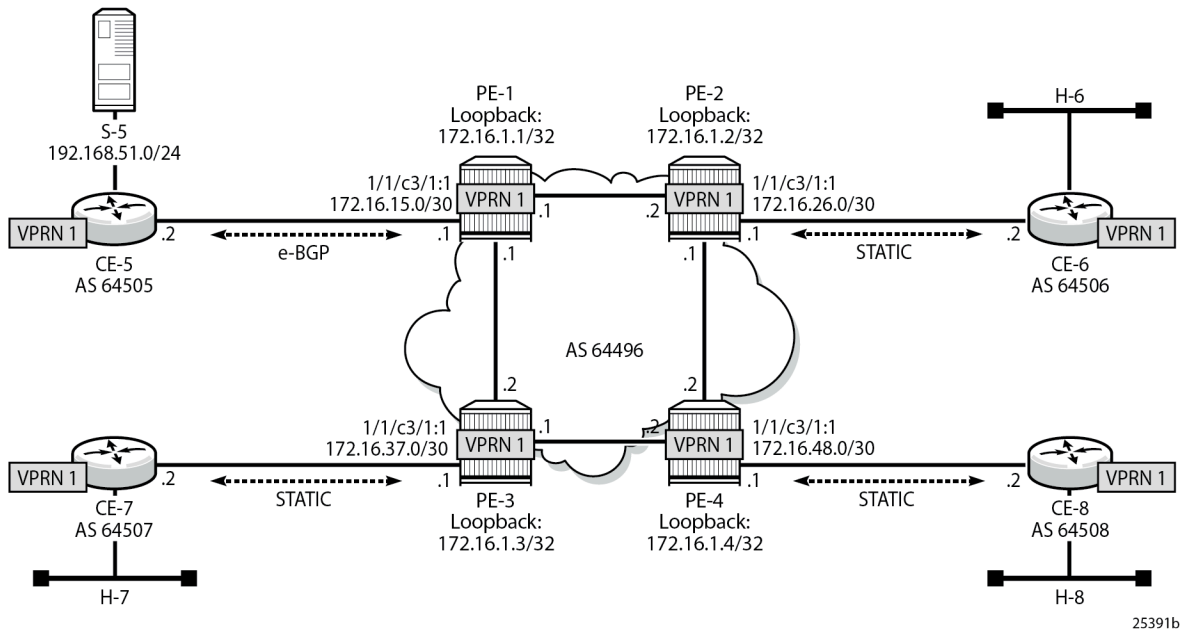
### PE VPRN Configuration and PE VPRN Multicast Configuration for NG-MVPN

A VPRN is created on each PE per service (the different services using mLDP, RSVP-TE, and AFI/SAFI with PIM); these are the multicast VPRNs. PE-1 is the PE containing the attachment circuit toward CE-5. CE-5 is the CE nearest to the source. PE-2, PE-3, and PE-4 contain attachment circuits toward CE-6, CE-7, and CE-8 respectively. Each CE has a receiving host attached.

### PMSI using mLDP

Figure 30: VPRN 1 Topology used for mLDP shows the details of the topology for VPRN 1.

Figure 30: VPRN 1 Topology used for mLDP





## Unicast

### Step 1.

Create VPRN 1 on each PE, containing a route-distinguisher of 64496:10X (where X= number of PE) and vrf-target of 64496:100. The autonomous system number is 64496. For the next hop tunnel route resolution to connect the VPRNs between the PEs, manually configured spoke SDPs are created (other methods such as auto-bind-tunnel resolution-filter LDP resolution filter could also be used). LDP was already enabled.

```
# on PE-1
configure
  service
    sdp 12 mpls create
      far-end 192.0.2.2
      ldp
      no shutdown
    exit
    sdp 13 mpls create
      far-end 192.0.2.3
      ldp
      no shutdown
    exit
    sdp 14 mpls create
      far-end 192.0.2.4
      ldp
      no shutdown
    exit
    vprn 1 name "1" customer 1 create
      description "mLDP"
      autonomous-system 64496
      bgp-ipvpn
        mpls
          route-distinguisher 64496:101
          vrf-target target:64496:100
          no shutdown
        exit
      exit
    spoke-sdp 12:1 create
    exit
    spoke-sdp 13:1 create
    exit
    spoke-sdp 14:1 create
    exit
    no shutdown
  exit
exit all
```

### Step 2.

Create an attachment circuit interface on PE-1 toward CE-5 and a loopback (the loopback is not mandatory, but it is configured to aid troubleshooting the routers).

```
# on PE-1
configure
  service
    vprn 1
      interface "loopback" create
        address 172.16.1.1/32
        loopback
```

```

exit
interface "int-PE-1-CE-5" create
  address 172.16.15.1/30
  sap 1/1/c3/1:1 create
  exit
exit
no shutdown
exit all

```

### Step 3.

The source address of the multicast stream needs to be reachable by all routers (PEs and CEs) within the VPN. This is advertised within BGP from CE-5 to PE-1. Create a BGP peering relationship with the CE as follows:

```

# on PE-1
configure
  service
    vprn 1
      bgp
        group "EXTERNAL"
          type external
          peer-as 64505
          neighbor 172.16.15.2
          exit
        exit
      no shutdown
    exit
  no shutdown
exit all

```

### Step 4.

On CE-5, create a VPRN to support the connection of the source to the CE and to connect the CE to the PE. Two attachment circuits are required as well as a BGP peering relationship with the PE. This uses a default BGP address family of ipv4.

```

# on CE-5
configure
  service
    vprn 1 customer 1 create
      autonomous-system 64505
      interface "int-CE-5-PE-1" create
        address 172.16.15.2/30
        sap 1/1/c1/1:1 create
        exit
      exit
      interface "int-CE-5-S-5" create
        address 192.168.51.1/24
        sap 1/1/c3/1 create
        exit
      exit
    bgp-ipvpn
      mpls
        route-distinguisher 64505:1
        no shutdown
      exit
    exit
  bgp
    group "EXTERNAL"
      type external
      peer-as 64496

```

```

neighbor 172.16.15.1
  exit
  exit
  no shutdown
exit
no shutdown
exit all

```

### Step 5.

For the subnet on the CE connecting to the source to be advertised within BGP, a route policy is required. The subnet containing the multicast source is 192.168.51.0/24, so a prefix-list can be used to define a match, and then used within a route policy to inject into BGP.

```

# on CE-5
configure
router
  policy-options
  begin
    prefix-list "SOURCE-PREFIX"
      prefix 192.168.51.0/24 exact
    exit
    policy-statement "EXPORT-SOURCE-PREFIX-T0-BGP-1"
      entry 10
        from
          prefix-list "SOURCE-PREFIX"
        exit
        to
          protocol bgp
        exit
        action accept
      exit
    exit
  commit
exit all

```

```

# on CE-5
configure
service
  vprn 1
  bgp
    export "EXPORT-SOURCE-PREFIX-T0-BGP-1"
  exit
exit all

```

### Step 6.

Check that the route is seen in PE-1:

```

*A:PE-1# show router 1 route-table 192.168.51.0/24
=====
Route Table (Service: 1)
=====
Dest Prefix[Flags]                               Type  Proto  Age           Pref
  Next Hop[Interface Name]                       Metric
-----
192.168.51.0/24                                  Remote BGP      00h01m19s  170
  172.16.15.2                                     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
=====
```

This prefix is also automatically advertised within the BGP VPRN to all other PEs, and is installed in VRF 1. For example, on PE-4, the source subnet 192.168.51.0/24 is received via BGP VPN with a next-hop of PE-1 (192.0.2.1):

```
*A:PE-4# show router 1 route-table 192.168.51.0/24

=====
Route Table (Service: 1)
=====
Dest Prefix[Flags]                Type   Proto   Age           Pref
Next Hop[Interface Name]         Metric
-----
192.168.51.0/24                   Remote BGP VPN 00h01m22s 170
192.0.2.1 (tunneled)              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
=====
```

Each CE containing a multicast receiver must be able to reach the source. As an example on CE-6, a static route is configured with next hop 172.16.26.1 of interface int-PE-2-CE-6.

```
# on CE-6
configure
service
  vprn 1 customer 1 create
  autonomous-system 64506
  interface "int-CE-6-H-6" create
  address 192.168.61.1/24
  sap 1/1/c2/1:1 create
  exit
  exit
  interface "int-CE-6-PE-2" create
  address 172.16.26.2/30
  sap 1/1/c1/1:1 create
  exit
  exit
  static-route-entry 192.168.51.0/24
  next-hop 172.16.26.1
  no shutdown
  exit
  exit
  bgp-ipvpn
  mpls
  route-distinguisher 64506:1
  no shutdown
  exit
  exit
  no shutdown
exit all
```

After **Steps 1 to 6**, all required unicast routing is provisioned. The following sections show the configuration of the multicast in the VPRN.

## Auto-Discovery and mLDP PMSI Establishment

The MP-BGP based auto-discovery is implemented with a dedicated address family defined in RFC 4760 MP\_REACH\_NLRI/MP\_UNREACH\_NLRI attributes, with AFI 1 (IPv4) or 2 (IPv6) SAFI 5 (temporary value assigned by IANA). This is the mechanism by which each PE advertises the presence of an MVPN to other PEs. This can be achieved using PIM (like in Rosen MVPN) or using BGP. With the default parameter, BGP is automatically chosen because the PMSIs are mLDP and PIM is not an option in this case. Any PE that is a member of an MVPN advertises to the other PEs using a BGP multi-protocol network layer reachability information (NLRI) update that is sent to all PEs within the AS. This update contains an Intra-AS I-PMSI auto-discovery route type, also known as an Intra-AD. These use an address family `mvpn-ipv4`, so each PE must be configured to originate and accept such updates (this was done earlier when configuring the families).

At this step (auto-discovery), the information about the PMSI is exchanged, but the PMSI is not instantiated.

As each PE contains a CE which is part of the multicast VRF, it is necessary to enable PIM on each interface containing the attachment circuit toward a CE, and to configure the I-PMSI multicast tunnel for the VRF. For the BGP routes to be accepted into the VRF, a route-target community is required (`vrf-target`). This is configured in the **configure service vprn 1 mvpn** context and, in this case is set to the same value as the unicast `vrf-target` (the `vrf-target` community in the **configure service vprn 1 bgp-ipvpn mpls vrf-target** context).

On each PE, the PIM and MVPN context within the VPRN instance are configured as follows:

```
# on PE-4
configure
  service
    vprn 1
      pim
        interface "loopback"
        exit
        interface "int-PE-4-CE-8"
        exit
        no shutdown
      exit
      mvpn
        auto-discovery default
        c-mcast-signaling bgp
        provider-tunnel
          inclusive
          mldp
          no shutdown
        exit
      exit
      vrf-target unicast
    exit
  exit
  no shutdown
exit all
```

When PIM SSM is used, the configuration always shows RP static with no RP entries (this is enabled by default when PIM is provisioned). For the BGP routes to be accepted into the VRF, a route-target

community is required (vrf-target). Although it is not mandatory for the mvpn target to be equal to the unicast target, Nokia recommends to use **vrf-target unicast** to avoid configuration mistakes and extra complexity.

The status of VPRN 1 on PE-1 is shown with the following output:

```
*A:PE-1# show router 1 mvpn

=====
MVPN 1 configuration data
=====
signaling           : Bgp                auto-discovery      : Default
UMH Selection        : Highest-Ip          SA withdrawn         : Disabled
intersite-shared     : Enabled              Persist SA           : Disabled
vrf-import           : N/A
vrf-export           : N/A
vrf-target           : unicast
C-Mcast Import RT   : target:192.0.2.1:2

ipmsi              : ldp
i-pmsi P2MP AdmSt  : Up
i-pmsi Tunnel Name : mpls-if-73728
Mdt-type             : sender-receiver
ipmsi UMH RM        : Disabled

BSR signalling       : none
Wildcard s-pmsi     : Disabled
Multistream-SPMSI   : Disabled
s-pmsi               : none
data-delay-interval : 3 seconds
enable-asm-mdt      : N/A
spmsi UMH RM        : Disabled

=====
```

The following shows a debug of an Intra-AD BGP update message received by PE-1 that was sent by PE-2. The message contains the PMSI tunnel type to be used (LDP P2MP LSP), LSP identification (root node, opaque value) and the type of BGP update (Type: Intra-AD Len: 12 RD: 64496:102 Orig: 192.0.2.2):

```
11 2023/10/25 09:17:14.574 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.2
"Peer 1: 192.0.2.2: UPDATE
Peer 1: 192.0.2.2 - Received BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 91
  Flag: 0x90 Type: 14 Len: 23 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.2
    Type: Intra-AD Len: 12 RD: 64496:102 Orig: 192.0.2.2
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:64496:100
  Flag: 0xc0 Type: 22 Len: 22 PMSI:
    Tunnel-type LDP P2MP LSP (2)
    Flags: (0x0)[Type: None BM: 0 U: 0 Leaf: not required]
    MPLS Label 0
    Root-Node 192.0.2.2, LSP-ID 0x2001
"
```

The setup has four PEs, so every PE should see the Intra-AD routes from its peers; the following output shows the routes received in PE-1:

```
*A:PE-1# show router bgp routes mvpn-ipv4 type intra-ad
=====
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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP MVPN-IPv4 Routes
=====
Flag  RouteType      OriginatorIP      LocalPref  MED
      RD           SourceAS          Path-Id     IGP Cost
      Nexthop      SourceIP          Label
      As-Path      GroupIP
-----
u*>i  Intra-Ad      192.0.2.2       100        0
      64496:102    -                None       -
      192.0.2.2    -
      No As-Path  -
u*>i  Intra-Ad      192.0.2.3       100        0
      64496:103    -                None       -
      192.0.2.3    -
      No As-Path  -
u*>i  Intra-Ad      192.0.2.4       100        0
      64496:104    -                None       -
      192.0.2.4    -
      No As-Path  -
-----
Routes : 3
=====
```

The detailed output of the Intra-AD received from PE-2 shows the Tunnel-Type LDP P2MP LSP (LSP-ID is 8193), the originator IP (192.0.2.2), and the route-distinguisher (64496:102):

```
*A:PE-1# show router bgp routes mvpn-ipv4 type intra-ad 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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP MVPN-IPv4 Routes
=====
Original Attributes

Route Type      : Intra-Ad
Route Dist.     : 64496:102
Originator IP   : 192.0.2.2
Nexthop        : 192.0.2.2
Path Id        : None
From           : 192.0.2.2
Res. Nexthop   : 0.0.0.0
Local Pref.    : 100
Aggregator AS  : None
Interface Name : NotAvailable
Aggregator     : None
```

```

Atomic Aggr.   : Not Atomic           MED           : 0
AIGP Metric   : None                 IGP Cost      : n/a
Connector     : None
Community     : no-export target:64496:100
Cluster       : No Cluster Members
Originator Id : None                 Peer Router Id : 192.0.2.2
Flags         : Used Valid Best IGP
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 : 00h01m55s
VPRN Imported : 1
    
```

```

-----
PMSI Tunnel Attributes :
Tunnel-type   : LDP P2MP LSP
Flags           : Type: RNVE(0) BM: 0 U: 0 Leaf: not required
MPLS Label     : 0
Root-Node     : 192.0.2.2           LSP-ID       : 8193
    
```

---snip---

```

-----
Routes : 3
=====
    
```

Because of the receiver-driven nature of mLDP, mLDP P2MP LSPs are set up unsolicited from the leaf PEs toward the head-end PE. The leaf PEs discover the head-end PE via I-PMSI/S-PMSI AD routes. The tunnel identifier carried in the PMSI attribute is used as the P2MP forwarding equivalence class (FEC) Element. The tunnel identifier consists of the address of the head-end PE, along with a generic LSP identifier value. The generic LSP identifier value is automatically generated by the head-end PE. The preceding show command displays the PMSI information with the detail of the root node (192.0.2.2) and the LSP-ID (8193). The PMSI was created after receiving the AD message from PE-2, where the following excerpt from the previous debug shows the same information (0x2001 in HEX is equal to 8193 in decimal).

```

Flag: 0xc0 Type: 22 Len: 22 PMSI:
      Tunnel-type LDP P2MP LSP (2)
      Flags: (0x0)[Type: None BM: 0 U: 0 Leaf: not required]
      MPLS Label 0
      Root-Node 192.0.2.2, LSP-ID 0x2001
    
```

After the mLDP P2MP LSPs are created, the I-PMSI is instantiated in the core:

```
*A:PE-1# show router 1 pim neighbor
```

```
=====
PIM Neighbor ipv4
=====
```

Interface Nbr Address	Nbr DR Prty	Up Time	Expiry Time	Hold Time
int-PE-1-CE-5 172.16.15.2	1	0d 00:01:42	0d 00:01:35	105
mpls-if-73729 192.0.2.2	1	0d 00:02:06	never	65535
mpls-if-73730 192.0.2.3	1	0d 00:01:59	never	65535
mpls-if-73731 192.0.2.4	1	0d 00:01:49	never	65535



```

-----
Neighbors : 4
=====

*A:PE-1# show router 1 pim tunnel-interface

=====
PIM Interfaces ipv4
=====
Interface                Originator Address  Adm  Opr  Transport Type
-----
mpls-if-73728             192.0.2.1           Up   Up   Tx-IPMSI
mpls-if-73729             192.0.2.2           Up   Up   Rx-IPMSI
mpls-if-73730             192.0.2.3           Up   Up   Rx-IPMSI
mpls-if-73731             192.0.2.4           Up   Up   Rx-IPMSI
-----
Interfaces : 4
=====

```

Every PE has created an I-PMSI to the other PEs. Checking the mLDP P2MP LSPs that are originated, transit, or destination to PE-1:

```

*A:PE-1# show router ldp bindings active p2mp ipv4

=====
LDP Bindings (IPv4 LSR ID 192.0.2.1)
(IPv6 LSR ID ::)
=====
Label Status:
  U - Label In Use, N - Label Not In Use, W - Label Withdrawn
  WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
  e - Label ELC
FEC Flags:
  LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
  BA - ASBR Backup FEC
=====
LDP Generic IPv4 P2MP Bindings (Active)
=====
P2MP-Id                Interface
RootAddr               Op
IngLbl                 EgrLbl
EgrNH                  EgrIf/LspId
-----
8193                   73728
192.0.2.1              Push
--                    524282
192.168.12.2          1/1/c1/1

8193                   73728
192.0.2.1              Push
--                    524282
192.168.13.2          1/1/c2/1

8193                   73729
192.0.2.2              Pop
524282                 --
--                    --

8193                   73729
192.0.2.2              Swap
524282                 524281
192.168.13.2          1/1/c2/1

```

```

8193                               73730
192.0.2.3                          Pop
524281                              --
--                                  --

8193                               73730
192.0.2.3                          Swap
524281                              524281
192.168.12.2                       1/1/c1/1

8193                               73731
192.0.2.4                          Pop
524280                              --
--                                  --

-----
No. of Generic IPv4 P2MP Active Bindings: 7
=====
---snip---

```

The two first entries in the output show the P2MP LSP where PE-1 is the root head-end (Push). The other two entries (Swap and Pop) correspond with transit and leaf for the P2MP LSPs originated by the other PEs. The command shows a P2MP-ID (8193) with an interface 73728 (matches with the **show router 1 pim tunnel interface** being the PIM interface created from PE-1) with two egress interfaces pointing to PE-2 and PE-3.

A similar command executed on PE-2 shows:

```

*A:PE-2# show router ldp bindings active p2mp ipv4

=====
LDP Bindings (IPv4 LSR ID 192.0.2.2)
          (IPv6 LSR ID ::)
=====
Label Status:
    U - Label In Use, N - Label Not In Use, W - Label Withdrawn
    WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
    e - Label ELC
FEC Flags:
    LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
    BA - ASBR Backup FEC
=====
LDP Generic IPv4 P2MP Bindings (Active)
=====
P2MP-Id          Interface
RootAddr         Op
IngLbl           EgrLbl
EgrNH            EgrIf/LspId
-----
8193              73729
192.0.2.1        Pop
524282           --
--              --

8193              73729
192.0.2.1        Swap
524282           524280
192.168.24.2    1/1/c1/1

---snip---
-----

```

```
No. of Generic IPv4 P2MP Active Bindings: 7
```

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

On PE-2, the first entry shows that PE-2 is a leaf of the P2MP LSP tree created by PE-1 (ingress label is 524282 which was the egress label to reach PE-2 and is popped). However, the second entry shows that PE-2 is transit for the P2MP LSP going to PE-4 (ingress label 524282, egress label 524280 next hop PE-4).

The same command on PE-4 shows:

```
*A:PE-4# show router ldp bindings active p2mp ipv4
```

```
-----  
LDP Bindings (IPv4 LSR ID 192.0.2.4)  
(IPv6 LSR ID ::)
```

```
-----  
Label Status:
```

```
U - Label In Use, N - Label Not In Use, W - Label Withdrawn  
WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route  
e - Label ELC
```

```
FEC Flags:
```

```
LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,  
BA - ASBR Backup FEC
```

```
-----  
LDP Generic IPv4 P2MP Bindings (Active)
```

```
-----  
P2MP-Id          Interface  
RootAddr         Op  
IngLbl           EgrLbl  
EgrNH            EgrIf/LspId  
-----  
8193             73731  
192.0.2.1        Pop  
524280           --  
--              --
```

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

```
-----  
No. of Generic IPv4 P2MP Active Bindings: 5
```

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

In the first entry, the root is PE-1 and the action is Pop, being the ingress label 524280, showing that this is another leaf for the P2MP LSP started on PE-1.

To complete the information, checking on PE-3, the first entry there is a Pop where the root is PE-1, and the ingress label is 524282:

```
*A:PE-3# show router ldp bindings active p2mp ipv4
```

```
-----  
LDP Bindings (IPv4 LSR ID 192.0.2.3)  
(IPv6 LSR ID ::)
```

```
-----  
Label Status:
```

```
U - Label In Use, N - Label Not In Use, W - Label Withdrawn  
WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route  
e - Label ELC
```

```
FEC Flags:
```

```
LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,  
BA - ASBR Backup FEC
```

```

=====
LDP Generic IPv4 P2MP Bindings (Active)
=====
P2MP-Id                Interface
RootAddr              0p
IngLbl                EgrLbl
EgrNH                 EgrIf/LspId
-----
8193                   73729
192.0.2.1             Pop
524282                --
--                    --

---snip---

-----
No. of Generic IPv4 P2MP Active Bindings: 5
=====
---snip---

```

As a summary, each root PE has a P2MP LSP with three leaves (the other PEs) and they are also transit points to the P2MP LSPs created in the other PEs. As an additional check, an OAM ping can show the different leaves that a P2MP LSP has:

```

*A:PE-1# oam p2mp-lsp-ping ldp 8193 sender-addr 192.0.2.1 detail
P2MP identifier 8193: | 88 bytes MPLS payload

=====
Leaf Information
=====
From          RTT          Return Code
-----
192.0.2.3     =1.60ms     EgressRtr(3)
192.0.2.2     =1.66ms     EgressRtr(3)
192.0.2.4     =2.72ms     EgressRtr(3)
=====

Total Leafs responded = 3
      round-trip min/avg/max = 1.60 / 1.99 / 2.72 ms

Responses based on return code:
EgressRtr(3)=3

```

An easy way to see the path that the LDP P2MP LSP follows for a specific leaf is the following command (such as LDP trace from PE-1 to PE-4):

```

*A:PE-1# oam ldp-treetrace prefix 192.0.2.4/32

ldp-treetrace for Prefix 192.0.2.4/32:

    192.168.24.2, ttl = 2 dst = 127.1.0.255 rc = EgressRtr status = Done
    Hops:      192.168.12.2

ldp-treetrace discovery state: Done
ldp-treetrace discovery status: ' OK '
Total number of discovered paths: 1
Total number of failed traces: 0

```

The command shows that on PE-4, there is an active leaf of the P2MP LSP, and that there is an intermediate hop on PE-2.

## Traffic Flow

The receiver H-8, connected to CE-8, wants to join the group 232.1.1.1 with source 192.168.51.1 and sends an IGMPv3 report toward CE-8. CE-8 recognizes the report and sends a PIM join toward the source, therefore, it reaches PE-1 where the source is connected to through CE-5. The following output shows the debug seen on PE-4, where the PIM join is received from CE-8 and a BGP update Source Join is sent to all PEs (only the update sent to PE-1 is shown).

```

17 2023/10/25 09:23:56.704 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimJPPProcessSG
pimJPPProcessSG: (S,G)-> (192.168.51.2,232.1.1.1) type <S,G>, i/f int-PE-4-CE-8,
upNbr 172.16.48.1 isJoin 1 isRpt 0 holdTime 210 isBierInband 0 srcAddr 172.16.48.2"

18 2023/10/25 09:23:56.705 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimRtmFindRpfNexthop
Track (192.168.51.2,232.1.1.1) type <S,G> using 192.168.51.2"

19 2023/10/25 09:23:56.705 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimRtmAddSrcEntry
Added src entry for src 192.168.51.2"

20 2023/10/25 09:23:56.705 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimJPPrintFsmEvent
PIM JP Downstream: State NoInfo Event RxJoin StandbyEvent F,
(S,G) (192.168.51.2,232.1.1.1) groupType <S,G>"

21 2023/10/25 09:23:56.705 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimJPPrintFsmEvent
PIM JP Upstream: State NotJoined Event JoinDesiredTrue StandbyEvent F,
(S,G) (192.168.51.2,232.1.1.1) groupType <S,G>"

22 2023/10/25 09:23:56.705 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimSGUpJoinDesiredTrue
No upstream interface. pSG (192.168.51.2,232.1.1.1) rpfType 3"

23 2023/10/25 09:23:56.705 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimSGUpJoinDesiredTrue
No upstream interface SG (192.168.51.2,232.1.1.1) rpfType 3"

24 2023/10/25 09:23:56.705 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimRtmProcessNhresEvent
RTM-Nhres Event U-RTM NEW Src 192.168.51.2 SrcRtmUse UCAST"

25 2023/10/25 09:23:56.705 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimRtmProcessNhresEvent
Prefix 192.168.51.0/24 numNextHops 1 owner BGP_VPN metric 0 pref 170"

26 2023/10/25 09:23:56.705 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimRtmSrcResolveSGsInt
Trying to resolve SG (192.168.51.2,232.1.1.1)"

27 2023/10/25 09:23:56.705 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimRtmNotifyRpfChange
RPF Change to Source/RP 192.168.51.2 for SG (192.168.51.2,232.1.1.1)
dynMLDP F via NH 192.0.2.1 IfIdx: 73731 RpfType: REMOTE Reason: RTE_ADD
old NH 0.0.0.0 IfIdx: 0 RpfType: NONE mplsRpf F NextHops 1 reg 1/1 lfa 0/0"

28 2023/10/25 09:23:56.705 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimRtmNotifyRpfChange
SG (192.168.51.2,232.1.1.1) Source/RP 192.168.51.2 Ipsid 73728 NhIf 0
new NhIf 73731"

```

```

29 2023/10/25 09:23:56.705 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimJPPrintFsmEvent
PIM JP Upstream: State Joined Event MribChange StandbyEvent F,
(S,G) (192.168.51.2,232.1.1.1) groupType <S,G>"

30 2023/10/25 09:23:56.705 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimSGUpStateJMribChange
SG (192.168.51.2,232.1.1.1), type <S,G> oldMribNhopIp 0.0.0.0
oldRpfNbrIp 0.0.0.0, oldRpfType NONE oldRpfIf 0 rptMribNhopIp 0.0.0.0,
rptRpfNbrIp 0.0.0.0 rtmReason 48 isSGExtNet : no"

31 2023/10/25 09:23:56.705 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimSGUpStateJMribChange
SG (192.168.51.2,232.1.1.1), type <S,G> newMribNhopIp 192.0.2.1
newRpfNbrIp 192.0.2.1 newRpfType REMOTE newRpfIf 73731"

32 2023/10/25 09:23:56.705 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimAddToJPTxPdu
pimAddToJPTxPdu: (S,G)-> (192.168.51.2,232.1.1.1), type <S,G>, txPendFlag J isStandby F"

33 2023/10/25 09:23:56.705 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimRtmUpdateSGMetric
SG metric 4294967295 pref 2147483647, new metric 0 pref 170"

---snip---

36 2023/10/25 09:23:56.705 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 76
  Flag: 0x90 Type: 14 Len: 33 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.4
    Type: Source-Join Len:22 RD: 64496:101 SrcAS: 64496 Src: 192.168.51.2 Grp: 232.1.1.1
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:192.0.2.1:2
"

```

The following debug shows that PE-1 receives the BGP update Source Join with source 192.168.1.1 and group 232.1.1.1 and sends a PIM join toward CE-5:

```

19 2023/10/25 09:23:55.781 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.4
"Peer 1: 192.0.2.4: UPDATE
Peer 1: 192.0.2.4 - Received BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 76
  Flag: 0x90 Type: 14 Len: 33 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.4
    Type: Source-Join Len:22 RD: 64496:101 SrcAS: 64496 Src: 192.168.51.2 Grp: 232.1.1.1
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:

```

```
no-export
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
  target:192.0.2.1:2
"
20 2023/10/25 09:23:55.781 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimProcessMvpnRouteMsg
originator 0.0.0.0: add rtType SOURCE_TREE_JOIN nextHop 192.0.2.4
source 192.168.51.2 group 232.1.1.1"
21 2023/10/25 09:23:55.781 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimJPPProcessSG
pimJPPProcessSG: (S,G)-> (192.168.51.2,232.1.1.1) type <S,G>, i/f mpls-if-73728,
upNbr 192.0.2.1 isJoin 1 isRpt 0 holdTime 65535 isBierInband 0 srcAddr 192.0.2.4"
22 2023/10/25 09:23:55.781 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimRtmFindRpfNextHop
Track (192.168.51.2,232.1.1.1) type <S,G> using 192.168.51.2"
23 2023/10/25 09:23:55.781 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimRtmAddSrcEntry
Added src entry for src 192.168.51.2"
24 2023/10/25 09:23:55.781 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimJPPrintFsmEvent
PIM JP Downstream: State NoInfo Event RxJoin StandbyEvent F,
(S,G) (192.168.51.2,232.1.1.1) groupType <S,G>"
25 2023/10/25 09:23:55.781 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimJPPrintFsmEvent
PIM JP Upstream: State NotJoined Event JoinDesiredTrue StandbyEvent F,
(S,G) (192.168.51.2,232.1.1.1) groupType <S,G>"
26 2023/10/25 09:23:55.781 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimAddToJPTxPdu
pimAddToJPTxPdu: (S,G)-> (192.168.51.2,232.1.1.1), type <S,G>, txPendFlag J isStandby F"
27 2023/10/25 09:23:55.781 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimRtmProcessNhresEvent
RTM-Nhres Event U-RTM NEW Src 192.168.51.2 SrcRtmUse UCAST"
28 2023/10/25 09:23:55.781 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimRtmProcessNhresEvent
Prefix 192.168.51.0/24 numNextHops 1 owner BGP metric 0 pref 170"
---snip---
37 2023/10/25 09:23:55.781 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimSGEncodeGroupSet
Encoding Join for source 192.168.51.2"
38 2023/10/25 09:23:55.781 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimSGEncodeGroupSet
num joined srcs 1, num pruned srcs 0"
39 2023/10/25 09:23:55.781 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimSendJoinPrunePdu
sending JP PDU with 1 groups, if 5 adj 172.16.15.2"
40 2023/10/25 09:23:55.781 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimSendJoinPrunePdu
if 5, adj 172.16.15.2. Nothing to send"
```

The BGP update source join received by PE-1 is displayed with the command:

```
*A:PE-1# show router bgp routes mvpn-ipv4 type source-join
=====
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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP MVPN-IPv4 Routes
=====
Flag RouteType      OriginatorIP      LocalPref  MED
      RD            SourceAS          Path-Id    IGP Cost
      Nexthop       SourceIP          Label
      As-Path       GroupIP
-----
u*>i  Source-Join      -                100        0
      64496:101        64496            None        -
      192.0.2.4      192.168.51.2
      No As-Path    232.1.1.1
-----
Routes : 1
=====
```

To verify the traffic: on PE-1 there is a group 232.1.1.1 with source 192.168.51.2, the Reverse Path Forwarding (RPF) is CE-5, the multicast traffic is flowing from CE-5 to PE-1 using int-PE-1-CE-5 and the outgoing interface is using the PMSI mLDP mpls-if-73728.

```
*A:PE-1# show router 1 pim group detail
=====
PIM Source Group ipv4
=====
Group Address      : 232.1.1.1
Source Address    : 192.168.51.2
RP Address           : 0
Advt Router         : 172.16.15.2
Flags               :                               Type           : (S,G)
Mode                : sparse
MRIB Next Hop       : 172.16.15.2
MRIB Src Flags      : remote
Keepalive Timer     : Not Running
Up Time             : 0d 00:00:47      Resolved By      : rtable-u

Up JP State         : Joined           Up JP Expiry      : 0d 00:00:13
Up JP Rpt           : Not Joined StarG  Up JP Rpt Override : 0d 00:00:00

Register State      : No Info
Reg From Anycast RP: No

Rpf Neighbor       : 172.16.15.2
Incoming Intf     : int-PE-1-CE-5
Outgoing Intf List: mpls-if-73728

Curr Fwding Rate    : 4815.728 kbps
Forwarded Packets   : 29050           Discarded Packets : 0
Forwarded Octets    : 28527100       RPF Mismatches    : 0
Spt threshold       : 0 kbps           ECMP opt threshold : 7
Admin bandwidth     : 1 kbps
```



```
-----
Groups : 1
=====
```

On PE-4, the same (S,G) arrives in the incoming interface mpls-if-73731 and the outgoing interface is int-PE-4-CE-8.

```
*A:PE-4# show router 1 pim group detail
```

```
=====
PIM Source Group ipv4
=====
Group Address      : 232.1.1.1
Source Address     : 192.168.51.2
RP Address         : 0
Advt Router        : 192.0.2.1
Flags              :
Mode               : sparse
MRIB Next Hop      : 192.0.2.1
MRIB Src Flags     : remote
Keepalive Timer    : Not Running
Up Time            : 0d 00:00:49
Resolved By        : rtable-u

Up JP State        : Joined
Up JP Rpt          : Not Joined StarG
Up JP Expiry       : 0d 00:00:11
Up JP Rpt Override : 0d 00:00:00

Register State     : No Info
Reg From Anycast RP: No

Rpf Neighbor      : 192.0.2.1
Incoming Intf     : mpls-if-73731
Outgoing Intf List : int-PE-4-CE-8

Curr Fwding Rate   : 4811.800 kbps
Forwarded Packets  : 30254
Forwarded Octets   : 29709428
Spt threshold      : 0 kbps
Admin bandwidth    : 1 kbps
Discarded Packets  : 0
RPF Mismatches     : 0
ECMP opt threshold : 7
-----
Groups : 1
=====
```

When the receiver is not interested in the channel group any more, the receiver H-8 sends an IGMPv3 leave, PE-4 sends a PIM prune translated to a BGP MP\_UNREACH NLRI to all PEs. As mentioned before, rapid withdrawals are sent without waiting for the MRAI (for simplicity, only one BGP update is shown in the output debug).

```
41 2023/10/25 09:25:11.972 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimJPPProcessSG
pimJPPProcessSG: (S,G)-> (192.168.51.2,232.1.1.1) type <S,G>, i/f int-PE-4-CE-8,
upNbr 172.16.48.1 isJoin 0 isRpt 0 holdTime 210 isBierInband 0 srcAddr 172.16.48.2"

42 2023/10/25 09:25:11.972 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimJPPrintFsmEvent
PIM JP Downstream: State Joined Event RxPrune StandbyEvent F,
(S,G) (192.168.51.2,232.1.1.1) groupType <S,G>"

43 2023/10/25 09:25:11.972 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimJPPrintFsmEvent
PIM JP Downstream: State PrunePending Event PrunePendTimerExp StandbyEvent F,
(S,G) (192.168.51.2,232.1.1.1) groupType <S,G>"
```

```

44 2023/10/25 09:25:11.972 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimJPPrintFsmEvent
PIM JP Upstream: State Joined Event JoinDesiredFalse StandbyEvent F,
(S,G) (192.168.51.2,232.1.1.1) groupType <S,G>"

45 2023/10/25 09:25:11.972 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimAddToJPTxPdu
pimAddToJPTxPdu: (S,G)-> (192.168.51.2,232.1.1.1), type <S,G>, txPendFlag P isStandby F"

46 2023/10/25 09:25:11.973 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimRtmStopRpfNextHop
Stop tracking (192.168.51.2,232.1.1.1) type <S,G> with 192.168.51.2 pRtmNhop 0x132a403e8"

47 2023/10/25 09:25:11.973 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 31
  Flag: 0x90 Type: 15 Len: 27 Multiprotocol Unreachable NLRI:
    Address Family MVPN_IPV4
    Type: Source-Join Len:22 RD: 64496:101 SrcAS: 64496 Src: 192.168.51.2 Grp: 232.1.1.1
"

---snip---

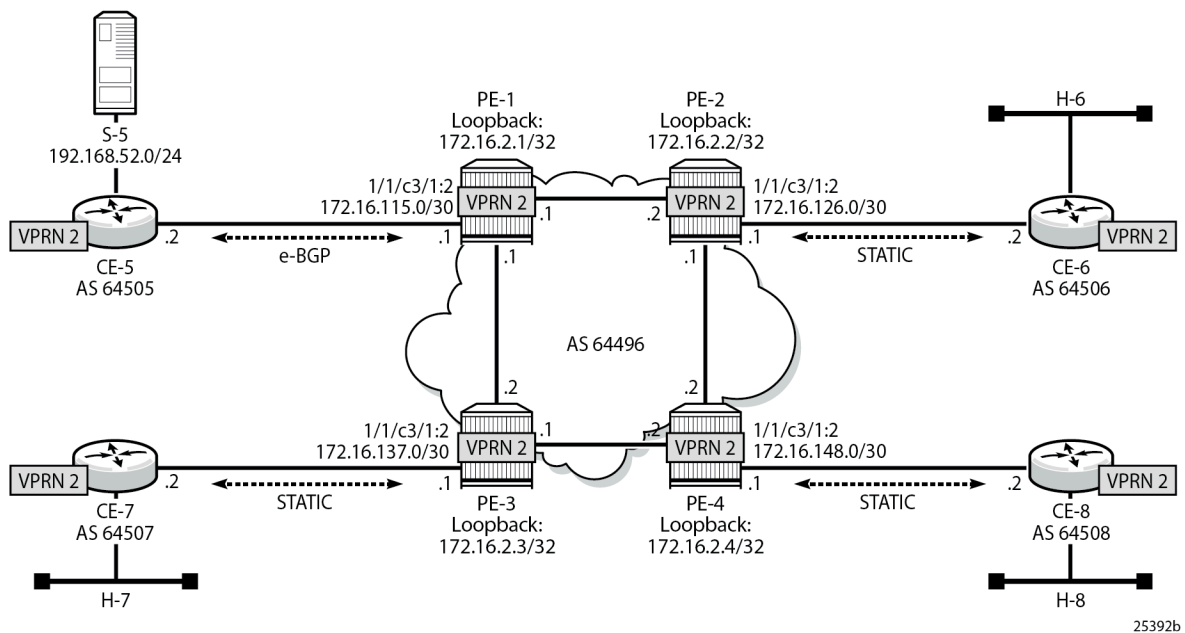
49 2023/10/25 09:25:11.973 UTC MINOR: DEBUG #2001 vprn1 PIM[vprn 2 vprn1]
"PIM[vprn 2 vprn1]: pimRtmDelSrcEntry
Deleted src entry for src 192.168.51.2"

```

## PMSI using RSVP-TE

Figure 31: VPRN 2 Topology used for RSVP-TE P2MP shows the details of the topology for VPRN 2.

Figure 31: VPRN 2 Topology used for RSVP-TE P2MP



## Unicast

For the sake of simplicity, check **Steps 1 to 6** in [PMSI using mLDP](#) for VPRN 2 creation information. The same steps are repeated for RSVP, check [Figure 31: VPRN 2 Topology used for RSVP-TE P2MP](#) for details. The result is the configuration in all the PEs, taking as an example PE-1:

```
# on PE-1
configure
  service
    vprn 2 customer 1 create
      description "P2MP RSVP"
      autonomous-system 64496
      interface "loopback" create
        address 172.16.2.1/32
        loopback
      exit
      interface "int-PE-1-CE-5" create
        address 172.16.115.1/30
        sap 1/1/c3/1:2 create
        exit
      exit
      bgp-ipvpn
        mpls
          route-distinguisher 64496:201
          vrf-target target:64496:200
          no shutdown
        exit
      exit
      bgp
        group "EXTERNAL"
          type external
          peer-as 64505
          neighbor 172.16.115.2
          exit
        exit
        no shutdown
      exit
      spoke-sdp 12:2 create
      exit
      spoke-sdp 13:2 create
      exit
      spoke-sdp 14:2 create
      exit
      no shutdown
    exit all
```

Because RSVP is the signaling protocol to establish the P2MP LSPs, RSVP is configured on the interfaces. In addition, to use P2MP RSVP, an LSP template is needed. The template defines the characteristics of the LSP to be created, for example, make-before-break, bandwidth, administrative groups, CSPF, specific paths, etc. A basic template is used here. TE parameters specified in the template are commonly used in each RSVP PATH message for each of the branches of the P2MP RSVP LSP. The template is used in the mvpn context within the VPRN configuration (see [Auto-Discovery and RSVP PMSI Establishment](#)). The resignal timer for P2MP is configured to the minimum value of sixty minutes (60 - 10080 minutes):

```
# on PE-1
configure
  router
    mpls
```

```

p2mp-resignal-timer 60
interface "system"
    no shutdown
exit
interface "int-PE-1-PE-2"
    no shutdown
exit
interface "int-PE-1-PE-3"
    no shutdown
exit
path "EMPTY"
    no shutdown
exit
lsp-template "VRF2" p2mp
    default-path "EMPTY"
    path-computation-method local-cspf
    fast-reroute facility
    exit
    no shutdown
exit
no shutdown
exit
    rsvp no shutdown
exit all

```

## Auto-Discovery and RSVP PMSI Establishment

The MP-BGP based auto-discovery is implemented with a new address family defined in RFC 4760 MP\_REACH\_NLRI/MP\_UNREACH\_NLRI attributes, with AFI 1 (IPv4) or 2 (IPv6) SAFI 5 (temporary value assigned by IANA). This is the mechanism by which each PE advertises the presence of an MVPN to other PEs. This can be achieved using PIM (like in Rosen MVPN) or using BGP. With the default parameter, BGP is automatically chosen because the PMSIs are RSVP and PIM is not an option in this case. Any PE that is a member of an MVPN advertises to the other PEs using a BGP multi-protocol network layer reachability information (NLRI) update that is sent to all PEs within the AS. This update contains an Intra-AS I-PMSI auto-discovery route type, also known as an Intra-AD. These use an address family mvpn-ipv4, so each PE must be configured to originate and accept such updates (this was done earlier when configuring the families).

At this step (auto-discovery), the information about the PMSI is exchanged, but the PMSI is not instantiated.

As each PE contains a CE which is part of the multicast VRF, it is necessary to enable PIM on each interface containing the attachment circuit toward a CE, and to configure the I-PMSI multicast tunnel for the VRF. For the BGP routes to be accepted into the VRF a route-target community is required (vrf-target). Although it is not mandatory for the MVPN vrf-target to be equal to the unicast target, Nokia recommends to use vrf-target unicast to avoid configuration mistakes and extra complexity.

On each PE, the multicast configuration in the VPRN instance is as follows:

```

# on PE-1
configure
    service
        vprn 2
            pim
                interface "loopback"
                exit
                interface "int-PE-1-CE-5"
                exit

```

```

no shutdown
exit
mvpn
  auto-discovery default
  c-mcast-signaling bgp
  provider-tunnel
    inclusive
    rsvp
      lsp-template "VRF2"
      no shutdown
    exit
  exit
exit
vrf-target unicast
exit
exit
exit all

```

The status of VPRN 2 on PE-1 is shown with the following output:

```

*A:PE-1# show router 2 mvpn
=====
MVPN 2 configuration data
=====
signaling           : Bgp          auto-discovery      : Default
UMH Selection       : Highest-Ip      SA withdrawn        : Disabled
intersite-shared    : Enabled          Persist SA          : Disabled
vrf-import          : N/A
vrf-export          : N/A
vrf-target          : unicast
C-Mcast Import RT  : target:192.0.2.1:3

ipmsi               : rsvp VRF2
i-pmsi P2MP AdmSt   : Up
i-pmsi Tunnel Name  : VRF2-2-73732
enable-bfd-root     : false           enable-bfd-leaf     : false
Mdt-type            : sender-receiver
ipmsi UMH RM        : Disabled

BSR signalling      : none
Wildcard s-pmsi    : Disabled
Multistream-SPMSI  : Disabled
s-pmsi              : none
data-delay-interval: 3 seconds
enable-asm-mdt     : N/A
spmsi UMH RM       : Disabled
=====

```

The following shows a debug of an Intra-AD BGP update message received by PE-1 that was sent by PE-4. The message contains the PMSI tunnel-type to be used (RSVP-TE P2MP LSP), the P2MP LSP ID (encoded as extended tunnel ID and P2MP-ID carried in the RSVP Session object), and the type of BGP update (Type: Intra-AD Len: 12 RD: 64496:204 Orig: 192.0.2.4):

```

29 2023/10/25 09:39:34.025 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.4
"Peer 1: 192.0.2.4: UPDATE
Peer 1: 192.0.2.4 - Received BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 86
  Flag: 0x90 Type: 14 Len: 23 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4

```

```

NextHop len 4 NextHop 192.0.2.4
Type: Intra-AD Len: 12 RD: 64496:204 Orig: 192.0.2.4
Flag: 0x40 Type: 1 Len: 1 Origin: 0
Flag: 0x40 Type: 2 Len: 0 AS Path:
Flag: 0x80 Type: 4 Len: 4 MED: 0
Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
Flag: 0xc0 Type: 8 Len: 4 Community:
no-export
Flag: 0xc0 Type: 16 Len: 8 Extended Community:
target:64496:200
Flag: 0xc0 Type: 22 Len: 17 PMSI:
Tunnel-type RSVP-TE P2MP LSP (1)
Flags: (0x0)[Type: None BM: 0 U: 0 Leaf: not required]
MPLS Label 0
P2MP-ID 0x2, Tunnel-ID: 61441, Extended-Tunnel-ID 192.0.2.4
"

```

The setup has four PEs, so every PE should see the others peer Intra-AD route; the following output shows the routes received in PE-1:

```

*A:PE-1# show router bgp routes mvpn-ipv4 type intra-ad
=====
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
Origin codes  : i - IGP, e - EGP, ? - incomplete

=====
BGP MVPN-IPv4 Routes
=====
Flag RouteType      OriginatorIP      LocalPref  MED
RD      SourceAS      Path-Id  IGP Cost
Nexthop SourceIP
As-Path GroupIP
-----
u*>i  Intra-Ad          192.0.2.2         100        0
      64496:202      -                 None       -
      192.0.2.2      -                 -          -
      No As-Path     -                 -          -
u*>i  Intra-Ad          192.0.2.3         100        0
      64496:203      -                 None       -
      192.0.2.3      -                 -          -
      No As-Path     -                 -          -
u*>i  Intra-Ad          192.0.2.4         100        0
      64496:204      -                 None       -
      192.0.2.4      -                 -          -
      No As-Path     -                 -          -
-----
Routes : 3
=====

```

The detailed output of the Intra-AD received from PE-4 shows the tunnel-type RSVP-TE P2MP LSP (P2MP-ID is 2), the originator id (192.0.2.4), and the route-distinguisher (64496:204):

```

*A:PE-1# show router bgp routes mvpn-ipv4 type intra-ad originator-ip 192.0.2.4 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

```

```

Origin codes   l - leaked, x - stale, > - best, b - backup, p - purge
               i - IGP, e - EGP, ? - incomplete

=====
BGP MVPN-IPv4 Routes
=====
Original Attributes

Route Type      : Intra-Ad
Route Dist.     : 64496:204
Originator IP   : 192.0.2.4
NextHop         : 192.0.2.4
Path Id         : None
From            : 192.0.2.4
Res. NextHop    : 0.0.0.0
Local Pref.     : 100
Aggregator AS   : None
Atomic Aggr.    : Not Atomic
AIGP Metric     : None
Connector       : None
Community       : no-export target:64496:200
Cluster         : No Cluster Members
Originator Id   : None
Flags           : Used Valid Best IGP
Route Source    : Internal
AS-Path         : No As-Path
Route Tag       : 0
Neighbor-AS     : n/a
DB Orig Val     : N/A
Source Class    : 0
Add Paths Send  : Default
Last Modified   : 00h01m34s
VPRN Imported   : 2

-----
PMSI Tunnel Attributes :
Tunnel-type       : RSVP-TE P2MP LSP
Flags             : Type: RNVE(0) BM: 0 U: 0 Leaf: not required
MPLS Label       : 0
P2MP-ID          : 2
Extended-Tunne*  : 192.0.2.4
Tunnel-ID        : 61441

-----

Modified Attributes

---snip---

-----
Routes : 1
=====
* indicates that the corresponding row element may have been truncated.

```

For the I-PMSI, the head-end PE firstly discovers all the leaf PEs via I-PMSI A-D routes, it then signals the P2MP LSP to all the leaf PEs using RSVP-TE (subsequently adding or removing S2L (source to leaf) paths as PEs are added or removed from the MVPN).

As in the mLDP case, the demarcation of the domains is in the PE. The PE router participates in both the customer multicast domain and the provider multicast domain. The customer CEs are limited to a multicast adjacency with the multicast instance on the PE created to support that specific customer IP-VPN. This way, customers are isolated from the provider core multicast domain and other customer multicast domains while the provider core P routers only participate in the provider multicast domain and are isolated from all customer multicast domains. C-trees to P-tunnels bindings are also discovered using BGP routes, instead of PIM join TLVs. MVPN c-multicast routing information is exchanged between PEs by using c-multicast routes that are carried using MCAST-VPN NLRIs.

After the RSVP-TE P2MP LSPs are created, the I-PMSI is instantiated in the core:

```
*A:PE-1# show router 2 pim neighbor
=====
PIM Neighbor ipv4
=====
Interface          Nbr DR Prty   Up Time      Expiry Time  Hold Time
-----
Nbr Address
-----
int-PE-1-CE-5      1             0d 00:01:34  0d 00:01:42  105
172.16.115.2
mpls-if-73733      1             0d 00:02:03  never        65535
192.0.2.2
mpls-if-73734      1             0d 00:01:55  never        65535
192.0.2.3
mpls-if-73735      1             0d 00:01:46  never        65535
192.0.2.4
-----
Neighbors : 4
=====
```

```
*A:PE-1# show router 2 pim tunnel-interface
=====
PIM Interfaces ipv4
=====
Interface          Originator Address  Adm  Opr  Transport Type
-----
mpls-if-73732      192.0.2.1           Up   Up   Tx-IPMSI
mpls-if-73733      192.0.2.2           Up   Up   Rx-IPMSI
mpls-if-73734      192.0.2.3           Up   Up   Rx-IPMSI
mpls-if-73735      192.0.2.4           Up   Up   Rx-IPMSI
-----
Interfaces : 4
=====
```

The following command displays the PMSIs created on a PE, taking PE-3 as an example:

```
*A:PE-3# show router 2 pim tunnel-interface
=====
PIM Interfaces ipv4
=====
Interface          Originator Address  Adm  Opr  Transport Type
-----
mpls-if-73732      192.0.2.3           Up   Up   Tx-IPMSI
mpls-if-73733      192.0.2.1           Up   Up   Rx-IPMSI
mpls-if-73734      192.0.2.2           Up   Up   Rx-IPMSI
mpls-if-73735      192.0.2.4           Up   Up   Rx-IPMSI
-----
Interfaces : 4
=====
```

```
*A:PE-3# tools dump router 2 mvpn provider-tunnels
=====
MVPN 2 Inclusive Provider Tunnels Originating
=====
ipmsi (RSVP)          P2MP-ID   Tunl-ID   Ext-Tunl-ID
-----
```



```

VRF2-2-73732                2                61441        192.0.2.3

=====
MVPN 2 Selective Provider Tunnels Originating
=====
spmsi (RSVP)                P2MP-ID      Tunl-ID      Ext-Tunl-ID
-----
No Tunnels Found
-----

=====
MVPN 2 Inclusive Provider Tunnels Terminating
=====
ipmsi (RSVP)                P2MP-ID      Tunl-ID      Ext-Tunl-ID
-----
mpls-if-73733                2             61441        192.0.2.1
mpls-if-73734                2             61441        192.0.2.2
mpls-if-73735                2             61441        192.0.2.4

=====
MVPN 2 Selective Provider Tunnels Terminating
=====
spmsi (RSVP)                P2MP-ID      Tunl-ID      Ext-Tunl-ID
-----
No Tunnels Found
-----

```

Every PE has created an I-PMSI to the other PEs. As an example, PE-1 has established an LSP with name VRF2-2-73732 with PE-2, PE-3 and PE-4 as leaves. The S2L path is empty because the template did not have any S2L path configured for simplicity.

```

*A:PE-1# show router mpls p2mp-lsp detail

=====
MPLS P2MP LSPs (Originating) (Detail)
=====
Legend :
+ - Inherited
=====
Type : Originating
-----
LSP Name      : VRF2-2-73732
LSP Type      : P2mpAutoLsp          LSP Tunnel ID      : 61441
LSP Index     : 61441              TTM Tunnel Id      : 61441
From          : 192.0.2.1
To            : Not configured
Adm State     : Up
LSP Up Time   : 0d 00:02:26
Transitions   : 1
Retry Limit   : 0
Signaling     : RSVP
Hop Limit     : 255
Adaptive      : Enabled
FastReroute   : Enabled
FR Method     : Facility
FR Node Protect : Disabled
FR Object     : Enabled
PathCompMethod : local-cspf
Oper State    : Up
LSP Down Time : 0d 00:00:00
Path Changes  : 1
Retry Timer   : 30 sec
Resv. Style   : SE
Negotiated MTU : n/a
ClassType     : 0
Oper FR       : Enabled
FR Hop Limit  : 16
FR Prop Adm Grp : Disabled
ADSPEC        : Disabled

```

```

FallbkPathComp : not-applicable
Metric          : Disabled
Load Bal Wt    : N/A
Include Grps   :
None
Least Fill     : Disabled
Metric Type    : igp
ClassForwarding : Disabled
Exclude Grps  :
None
Soft Preemption : Disabled

Revert Timer   : Disabled
Auto BW        : Disabled
LdpOverRsvp   : Enabled
VprnAutoBind  : Enabled
IGP Shortcut   : Enabled
IGP LFA        : Disabled
AllowSrOverSrte : Disabled
BGPTransTun   : Enabled
Oper Metric    : Disabled
Prop Adm Grp  : Disabled

P2MPInstance  : 2
P2MP-Inst-type : Primary
S2L Cfg Counter : 3
S2L Oper Counter : 3
S2L-Name      : EMPTY
To            : 192.0.2.2
S2L-Name      : EMPTY
To            : 192.0.2.3
S2L-Name      : EMPTY
To            : 192.0.2.4
=====

```

Checking the RSVP-TE P2MP LSPs that are originated, transit, or destination to PE-1, the show command allows filtering by type, in this case showing the originated LSPs only:

```

*A:PE-1# show router mpls p2mp-info type originate
=====
MPLS P2MP LSPs (Originate)
=====
-----
S2L:VRF2-2-73732::EMPTY
-----
Source IP Address : 192.0.2.1      Tunnel ID   : 61441
P2MP ID           : 2             Lsp ID     : 37888
To                : 192.0.2.2
Out Interface     : 1/1/c1/1      Out Label  : 524282
Num. of S2ls     : 1
-----
S2L:VRF2-2-73732::EMPTY
-----
Source IP Address : 192.0.2.1      Tunnel ID   : 61441
P2MP ID           : 2             Lsp ID     : 37888
To                : 192.0.2.3
Out Interface     : 1/1/c2/1      Out Label  : 524280
Num. of S2ls     : 2
-----
S2L:VRF2-2-73732::EMPTY
-----
Source IP Address : 192.0.2.1      Tunnel ID   : 61441
P2MP ID           : 2             Lsp ID     : 37888
To              : 192.0.2.4
Out Interface     : 1/1/c2/1      Out Label  : 524280
Num. of S2ls     : 2
-----
P2MP Cross-connect instances : 3

```

Following the path of the S2L from PE-1 to PE-4 (third entry S2L VRF2-2-73732), the outgoing interface is 1/1/c2/1 that connects PE-1 to PE-3, so the LSP goes to PE-4 via PE-3. The return path need not be via PE-3; it may be via PE-2.

```
*A:PE-3# show router mpls p2mp-info type transit
```

```
=====
MPLS P2MP LSPs (Transit)
=====
-----
S2L:VRF2-2-73732::EMPTY
-----
Source IP Address   : 192.0.2.1           Tunnel ID   : 61441
P2MP ID             : 2                 Lsp ID     : 37888
To                  : 192.0.2.4
Out Interface       : 1/1/c2/1         Out Label  : 524276
Num. of S2ls       : 1
-----
P2MP Cross-connect instances : 1
=====
```

As transit, PE-3 shows that there is an LSP coming from PE-1 (VRF2-2-73732) and the outgoing interface is 1/1/c2/1 that connects PE-3 with PE-4.

Using the same command with a different filter on PE-4, 3 P2MP LSPs are terminated, one from each remote PE (PE-1, PE-2 and PE-3). On PE-4, an S2L VRF2-2-73732 from 192.0.2.1 and P2MP ID = 2 is traced.

```
*A:PE-4# show router mpls p2mp-info type terminate
```

```
=====
MPLS P2MP LSPs (Terminate)
=====
-----
S2L:VRF2-2-73732::EMPTY
-----
Source IP Address   : 192.0.2.1           Tunnel ID   : 61441
P2MP ID             : 2                 Lsp ID     : 37888
To                  : 192.0.2.4
In Interface        : 1/1/c1/1         In Label   : 524276
Num. of S2ls       : 1
-----
S2L:VRF2-2-73732::EMPTY
-----
Source IP Address   : 192.0.2.2           Tunnel ID   : 61441
P2MP ID             : 2                 Lsp ID     : 12288
To                  : 192.0.2.4
In Interface        : 1/1/c2/1         In Label   : 524278
Num. of S2ls       : 1
-----
S2L:VRF2-2-73732::EMPTY
-----
Source IP Address   : 192.0.2.3           Tunnel ID   : 61441
P2MP ID             : 2                 Lsp ID     : 21504
To                  : 192.0.2.4
In Interface        : 1/1/c1/1         In Label   : 524277
Num. of S2ls       : 1
-----
P2MP Cross-connect instances : 3
=====
```

The following output shows the P2MP LSP on PE-1 with more detail:

```
*A:PE-1# show router mpls p2mp-lsp "VRF2-2-73732" p2mp-instance "2" s2l "EMPTY" detail
```

```
=====
MPLS LSP VRF2-2-73732 S2L EMPTY (Detail)
=====
```

Legend :

```

@ - Detour Available          # - Detour In Use
b - Bandwidth Protected      n - Node Protected
S - Strict                   L - Loose
A - ABR
s - Soft Preemption

```

```
=====
LSP Name       : VRF2-2-73732
S2L LSP ID    : 37888
P2MP ID       : 2                S2L Grp Id       : 1
Admin State   : Up              Oper State      : Up
S2L State:    : Active          :
S2L Name      : EMPTY
To            : 192.0.2.2
S2L Admin     : Up              S2L Oper       : Up
OutInterface  : 1/1/c1/1        Out Label      : 524282
S2L Up Time   : 0d 00:03:18    S2L Dn Time   : 0d 00:00:00
RetryAttempt  : 0              NextRetryIn   : 0 sec
S2L Trans     : 1              CSPF Queries  : 1
Failure Code  : noError        Failure Node   : n/a
Inter-area    : False
ExplicitHops  :
  No Hops Specified
Actual Hops   :
  192.168.12.1(192.0.2.1) @    Record Label   : N/A
  -> 192.168.12.2(192.0.2.2)  Record Label   : 524282
ComputedHops  :
  192.168.12.1(S)
  -> 192.168.12.2(S)
LastResignal  : n/a
=====
```

```
-----
LSP Name       : VRF2-2-73732
S2L LSP ID    : 37888
P2MP ID       : 2                S2L Grp Id       : 2
Admin State   : Up              Oper State      : Up
S2L State:    : Active          :
S2L Name      : EMPTY
To            : 192.0.2.3
S2L Admin     : Up              S2L Oper       : Up
OutInterface  : 1/1/c2/1        Out Label      : 524280
S2L Up Time   : 0d 00:03:09    S2L Dn Time   : 0d 00:00:00
RetryAttempt  : 0              NextRetryIn   : 0 sec
S2L Trans     : 1              CSPF Queries  : 1
Failure Code  : noError        Failure Node   : n/a
Inter-area    : False
ExplicitHops  :
  No Hops Specified
Actual Hops   :
  192.168.13.1(192.0.2.1) @    Record Label   : N/A
  -> 192.168.13.2(192.0.2.3)  Record Label   : 524280
ComputedHops  :
  192.168.13.1(S)
  -> 192.168.13.2(S)
LastResignal  : n/a
=====
```

```

-----
LSP Name       : VRF2-2-73732
S2L LSP ID    : 37888
P2MP ID       : 2                S2L Grp Id       : 3
Admin State   : Up                Oper State     : Up
S2L State:    : Active           :
S2L Name      : EMPTY
To            : 192.0.2.4
S2L Admin     : Up                S2L Oper       : Up
OutInterface  : 1/1/c2/1          Out Label      : 524280
S2L Up Time   : 0d 00:03:00      S2L Dn Time    : 0d 00:00:00
RetryAttempt  : 0                NextRetryIn    : 0 sec
S2L Trans     : 1                CSPF Queries   : 1
Failure Code  : noError          Failure Node    : n/a
Inter-area    : False
ExplicitHops  :
  No Hops Specified
Actual Hops   :
  192.168.13.1(192.0.2.1) @      Record Label   : N/A
  -> 192.168.13.2(192.0.2.3) @   Record Label   : 524280
  -> 192.168.34.2(192.0.2.4)    Record Label   : 524276
ComputedHops  :
  192.168.13.1(S)
  -> 192.168.13.2(S)
  -> 192.168.34.2(S)
LastResignal  : n/a
=====

```

The last entry, VRF2-2-73732, provides the details of the S2L traced earlier, displaying the different hops (PE-1, PE-3, and PE-4), the fast reroute protection (link protection is supported only) and the labels used (524280 from PE-1 to PE-3, 524276 from PE-3 to PE-4). On PE-1, although only one has been shown, both links PE-1 to PE-3 and PE-1 to PE-2 are fast reroute protected.

If any of the protected links between PE-1 and PE-2 or PE-3 are broken, fast reroute is initiated. The protected bypass hops are displayed with the following command:

```
*A:PE-1# show router mpls bypass-tunnel protected-lsp p2mp detail
```

```
=====
MPLS Bypass Tunnels (Detail)
=====
```

```
-----
bypass-link192.168.12.2-61442
-----
```

```

-----
To           : 192.168.24.1        State          : Up
Out I/F      : 1/1/c2/1          Out Label      : 524282
Up Time      : 0d 00:03:18       Active Time    : n/a
Reserved BW  : 0 Kbps            Protected LSP Count : 2
Type         : P2mp              Bypass Path Cost : 30
Setup Priority : 7                Hold Priority   : 0
Class Type   : 0
Exclude Node : None              Inter-Area     : False
Computed Hops :
  192.168.13.1(S)                Egress Admin Groups : None
  -> 192.168.13.2(S)              Egress Admin Groups : None
  -> 192.168.34.2(S)              Egress Admin Groups : None
  -> 192.168.24.1(S)              Egress Admin Groups : None
Actual Hops  :
  192.168.13.1(192.0.2.1)        Record Label   : N/A
  -> 192.168.13.2(192.0.2.3)      Record Label   : 524282
  -> 192.168.34.2(192.0.2.4)      Record Label   : 524282
  -> 192.168.24.1(192.0.2.2)      Record Label   : 524281

```

```
Protected LSPs -
LSP Name      : VRF2-2-73732::EMPTY
From          : 192.0.2.1           To          : 192.0.2.2
Avoid Node/Hop : 192.168.12.2       Downstream Label : 524282
Bandwidth     : 0 Kbps

LSP Name      : VRF2-2-73732::EMPTY
From          : 192.0.2.3           To          : 192.0.2.2
Avoid Node/Hop : 192.168.12.2       Downstream Label : 524279
Bandwidth     : 0 Kbps
```

```
-----
bypass-link192.168.13.2-61443
-----
```

```
To          : 192.168.34.1           State       : Up
Out I/F     : 1/1/c1/1             Out Label   : 524280
Up Time     : 0d 00:03:09          Active Time : n/a
Reserved BW : 0 Kbps              Protected LSP Count : 3
Type       : P2mp                 Bypass Path Cost : 30
Setup Priority : 7                 Hold Priority  : 0
Class Type  : 0
Exclude Node : None              Inter-Area    : False
Computed Hops :
  192.168.12.1(S)                Egress Admin Groups : None
  -> 192.168.12.2(S)             Egress Admin Groups : None
  -> 192.168.24.2(S)             Egress Admin Groups : None
  -> 192.168.34.1(S)             Egress Admin Groups : None
Actual Hops :
  192.168.12.1(192.0.2.1)        Record Label : N/A
  -> 192.168.12.2(192.0.2.2)    Record Label   : 524280
  -> 192.168.24.2(192.0.2.4)    Record Label   : 524280
  -> 192.168.34.1(192.0.2.3)    Record Label   : 524278
```

```
Protected LSPs -
LSP Name      : VRF2-2-73732::EMPTY
From          : 192.0.2.1           To          : 192.0.2.3
Avoid Node/Hop : 192.168.13.2       Downstream Label : 524280
Bandwidth     : 0 Kbps

LSP Name      : VRF2-2-73732::EMPTY
From          : 192.0.2.2           To          : 192.0.2.3
Avoid Node/Hop : 192.168.13.2       Downstream Label : 524279
Bandwidth     : 0 Kbps

LSP Name      : VRF2-2-73732::EMPTY
From          : 192.0.2.1           To          : 192.0.2.4
Avoid Node/Hop : 192.168.13.2       Downstream Label : 524280
Bandwidth     : 0 Kbps
```

## Traffic Flow

The receiver H-8, connected to CE-8, wants to join the group 232.2.2.2 with source 192.168.52.1 and so sends an IGMPv3 report toward CE-8. CE-8 recognizes the report and sends a PIM join toward the source, therefore, it reaches PE-1 where the source is connected to through CE-5. The following output shows the debug seen on PE-4, where the PIM join is received from CE-8 and a BGP update Source Join is sent to all PEs (only the update sent to PE-1 is shown).

```
1 2023/10/25 09:44:23.307 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
```

```
"PIM[vprn 3 vprn2]: pimJPPProcessSG
pimJPPProcessSG: (S,G)-> (192.168.52.2,232.2.2.2) type <S,G>, i/f int-PE-4-CE-8,
upNbr 172.16.148.1 isJoin 1 isRpt 0 holdTime 210 isBierInband 0 srcAddr 172.16.148.2"

2 2023/10/25 09:44:23.307 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimRtmFindRpfNextHop
Track (192.168.52.2,232.2.2.2) type <S,G> using 192.168.52.2"

3 2023/10/25 09:44:23.307 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimRtmAddSrcEntry
Added src entry for src 192.168.52.2"

4 2023/10/25 09:44:23.307 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimJPPrintFsmEvent
PIM JP Downstream: State NoInfo Event RxJoin StandbyEvent F,
(S,G) (192.168.52.2,232.2.2.2) groupType <S,G>"

5 2023/10/25 09:44:23.307 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimJPPrintFsmEvent
PIM JP Upstream: State NotJoined Event JoinDesiredTrue StandbyEvent F,
(S,G) (192.168.52.2,232.2.2.2) groupType <S,G>"

6 2023/10/25 09:44:23.307 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimSGUpJoinDesiredTrue
No upstream interface. pSG (192.168.52.2,232.2.2.2) rpfType 3"

7 2023/10/25 09:44:23.307 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimSGUpJoinDesiredTrue
No upstream interface SG (192.168.52.2,232.2.2.2) rpfType 3"

8 2023/10/25 09:44:23.307 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimRtmProcessNhresEvent
RTM-Nhres Event U-RTM NEW Src 192.168.52.2 SrcRtmUse UCAST"

9 2023/10/25 09:44:23.308 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimRtmProcessNhresEvent
Prefix 192.168.52.0/24 numNextHops 1 owner BGP_VPN metric 0 pref 170"

10 2023/10/25 09:44:23.308 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimRtmSrcResolveSGsInt
Trying to resolve SG (192.168.52.2,232.2.2.2)"

11 2023/10/25 09:44:23.308 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimRtmNotifyRpfChange
RPF Change to Source/RP 192.168.52.2 for SG (192.168.52.2,232.2.2.2)
dynMLDP F via NH 192.0.2.1 IfIdx: 73735 RpfType: REMOTE Reason: RTE_ADD
old NH 0.0.0.0 IfIdx: 0 RpfType: NONE mplsRpf F NextHops 1 reg 1/1 lfa 0/0"

12 2023/10/25 09:44:23.308 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimRtmNotifyRpfChange
SG (192.168.52.2,232.2.2.2) Source/RP 192.168.52.2 Ipmsi 73732 NhIf 0
new NhIf 73735"

13 2023/10/25 09:44:23.308 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimJPPrintFsmEvent
PIM JP Upstream: State Joined Event MribChange StandbyEvent F,
(S,G) (192.168.52.2,232.2.2.2) groupType <S,G>"

14 2023/10/25 09:44:23.308 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimSGUpStateJMribChange
SG (192.168.52.2,232.2.2.2), type <S,G> oldMribNhopIp 0.0.0.0
oldRpfNbrIp 0.0.0.0, oldRpfType NONE oldRpfIf 0 rptMribNhopIp 0.0.0.0,
rptRpfNbrIp 0.0.0.0 rtmReason 48 isSGExtNet : no"
```

```

15 2023/10/25 09:44:23.308 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimSGUpStateJMribChange
SG (192.168.52.2,232.2.2.2), type <S,G> newMribNhopIp 192.0.2.1
newRpfNbrIp 192.0.2.1 newRpfType REMOTE newRpfIf 73735"

16 2023/10/25 09:44:23.308 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimAddToJPTxPdu
pimAddToJPTxPdu: (S,G)-> (192.168.52.2,232.2.2.2), type <S,G>, txPendFlag J isStandby F"

17 2023/10/25 09:44:23.308 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimRtmUpdateSGMetric
SG metric 4294967295 pref 2147483647, new metric 0 pref 170"

---snip---

20 2023/10/25 09:44:23.308 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 76
  Flag: 0x90 Type: 14 Len: 33 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.4
    Type: Source-Join Len:22 RD: 64496:201 SrcAS: 64496 Src: 192.168.52.2 Grp: 232.2.2.2
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:192.0.2.1:3
"
```

The following debug shows that PE-1 receives the BGP update Source Join with source 192.168.52.1 and group 232.2.2.2 and sends a PIM join toward CE-5:

```

1 2023/10/25 09:44:22.386 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.4
"Peer 1: 192.0.2.4: UPDATE
Peer 1: 192.0.2.4 - Received BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 76
  Flag: 0x90 Type: 14 Len: 33 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.4
    Type: Source-Join Len:22 RD: 64496:201 SrcAS: 64496 Src: 192.168.52.2 Grp: 232.2.2.2
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:192.0.2.1:3
"
```

```

2 2023/10/25 09:44:22.386 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimProcessMvpnRouteMsg
originator 0.0.0.0: add rtType SOURCE_TREE_JOIN nextHop 192.0.2.4
source 192.168.52.2 group 232.2.2.2"

3 2023/10/25 09:44:22.386 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimJPPProcessSG
```



```
pimJPPProcessSG: (S,G)-> (192.168.52.2,232.2.2.2) type <S,G>, i/f mpls-if-73732,
upNbr 192.0.2.1 isJoin 1 isRpt 0 holdTime 65535 isBierInband 0 srcAddr 192.0.2.4"

4 2023/10/25 09:44:22.386 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimRtmFindRpfNexthop
Track (192.168.52.2,232.2.2.2) type <S,G> using 192.168.52.2"

5 2023/10/25 09:44:22.386 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimRtmAddSrcEntry
Added src entry for src 192.168.52.2"

6 2023/10/25 09:44:22.386 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimJPPrintFsmEvent
PIM JP Downstream: State NoInfo Event RxJoin StandbyEvent F,
(S,G) (192.168.52.2,232.2.2.2) groupType <S,G>"

7 2023/10/25 09:44:22.386 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimJPPrintFsmEvent
PIM JP Upstream: State NotJoined Event JoinDesiredTrue StandbyEvent F,
(S,G) (192.168.52.2,232.2.2.2) groupType <S,G>"

8 2023/10/25 09:44:22.386 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimAddToJPTxPdu
pimAddToJPTxPdu: (S,G)-> (192.168.52.2,232.2.2.2), type <S,G>, txPendFlag J isStandby F"

9 2023/10/25 09:44:22.386 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimRtmProcessNhresEvent
RTM-Nhres Event U-RTM NEW Src 192.168.52.2 SrcRtmUse UCAST"

10 2023/10/25 09:44:22.386 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimRtmProcessNhresEvent
Prefix 192.168.52.0/24 numNextHops 1 owner BGP metric 0 pref 170"

---snip---

19 2023/10/25 09:44:22.386 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimSGEncodeGroupSet
Encoding Join for source 192.168.52.2"

20 2023/10/25 09:44:22.386 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimSGEncodeGroupSet
num joined srcs 1, num pruned srcs 0"

21 2023/10/25 09:44:22.386 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimSendJoinPrunePdu
sending JP PDU with 1 groups, if 7 adj 172.16.115.2"
```

The BGP update source join received by PE-1 is displayed with the following command:

```
*A:PE-1# show router bgp routes mvpn-ipv4 type source-join
=====
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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP MVPN-IPv4 Routes
=====
Flag  RouteType      OriginatorIP      LocalPref  MED
      RD            SourceAS          Path-Id     IGP Cost
```

	Nexthop As-Path	SourceIP GroupIP		Label
u*>i	<b>Source-Join</b>	-	100	0
	64496:201	64496	None	-
	<b>192.0.2.4</b>	<b>192.168.52.2</b>		
	<b>No As-Path</b>	<b>232.2.2.2</b>		
-----				
Routes : 1				
=====				

To verify the traffic: on PE-1, there is a group 232.2.2.2 with source 192.168.52.1, the RPF is CE-5, and the multicast traffic is flowing from CE-5 to PE-1 using int-PE-1-CE-5 and the outgoing interface is using the PMSI RSVP mpls-if-73732.

```
*A:PE-1# show router 2 pim group detail
=====
PIM Source Group ipv4
=====
Group Address      : 232.2.2.2
Source Address    : 192.168.52.2
RP Address          : 0
Advt Router        : 172.16.115.2
Flags              :
Mode               : sparse
MRIB Next Hop     : 172.16.115.2
MRIB Src Flags    : remote
Keepalive Timer   : Not Running
Up Time           : 0d 00:00:40
Resolved By       : rtable-u

Up JP State       : Joined
Up JP Rpt         : Not Joined StarG
Up JP Expiry      : 0d 00:00:19
Up JP Rpt Override : 0d 00:00:00

Register State    : No Info
Reg From Anycast RP: No

Rpf Neighbor      : 172.16.115.2
Incoming Intf    : int-PE-1-CE-5
Outgoing Intf List : mpls-if-73732

Curr Fwding Rate  : 4811.800 kbps
Forwarded Packets : 24393
Forwarded Octets  : 23953926
Spt threshold     : 0 kbps
Admin bandwidth   : 1 kbps
Discarded Packets : 0
RPF Mismatches    : 0
ECMP opt threshold : 7
-----
Groups : 1
=====
```

On PE-4, the same (S,G) arrives in the incoming interface mpls-if-73735 and the outgoing interface is int-PE-4-CE-8.

```
*A:PE-4# show router 2 pim group detail
=====
PIM Source Group ipv4
=====
Group Address      : 232.2.2.2
Source Address     : 192.168.52.2
RP Address         : 0
Advt Router        : 192.0.2.1
```

```

Flags          :                               Type          : (S,G)
Mode           : sparse
MRIB Next Hop : 192.0.2.1
MRIB Src Flags : remote
Keepalive Timer : Not Running
Up Time        : 0d 00:00:43      Resolved By    : rtable-u

Up JP State    : Joined           Up JP Expiry   : 0d 00:00:16
Up JP Rpt      : Not Joined StarG Up JP Rpt Override : 0d 00:00:00

Register State : No Info
Reg From Anycast RP: No

Rpf Neighbor      : 192.0.2.1
Incoming Intf     : mpls-if-73735
Outgoing Intf List : int-PE-4-CE-8

Curr Fwding Rate : 4811.800 kbps
Forwarded Packets : 26211           Discarded Packets : 0
Forwarded Octets  : 25739202       RPF Mismatches    : 0
Spt threshold     : 0 kbps          ECMP opt threshold : 7
Admin bandwidth   : 1 kbps

```

-----  
Groups : 1  
=====

When the receiver is not interested in the channel group anymore, the receiver H-8 sends an IGMPv3 leave, PE-4 sends a PIM prune translated to a BGP MP\_UNREACH NLRI to all PEs. As mentioned before, rapid withdrawals are sent without waiting for the MRAL (for simplicity, only one BGP update is shown in the output debug).

```

21 2023/10/25 09:51:48.972 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimJPPProcessSG
pimJPPProcessSG: (S,G)-> (192.168.52.2,232.2.2.2) type <S,G>, i/f int-PE-4-CE-8,
upNbr 172.16.148.1 isJoin 0 isRpt 0 holdTime 210 isBierInband 0 srcAddr 172.16.148.2"

22 2023/10/25 09:51:48.972 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimJPPrintFsmEvent
PIM JP Downstream: State Joined Event RxPrune StandbyEvent F,
(S,G) (192.168.52.2,232.2.2.2) groupType <S,G>"

23 2023/10/25 09:51:48.972 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimJPPrintFsmEvent
PIM JP Downstream: State PrunePending Event PrunePendTimerExp StandbyEvent F,
(S,G) (192.168.52.2,232.2.2.2) groupType <S,G>"

24 2023/10/25 09:51:48.972 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimJPPrintFsmEvent
PIM JP Upstream: State Joined Event JoinDesiredFalse StandbyEvent F,
(S,G) (192.168.52.2,232.2.2.2) groupType <S,G>"

25 2023/10/25 09:51:48.972 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimAddToJPTxPdu
pimAddToJPTxPdu: (S,G)-> (192.168.52.2,232.2.2.2), type <S,G>, txPendFlag P isStandby F"

26 2023/10/25 09:51:48.972 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimRtmStopRpfNextHop
Stop tracking (192.168.52.2,232.2.2.2) type <S,G> with 192.168.52.2 pRtmNhop 0x132a402b8"

27 2023/10/25 09:51:48.973 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimRtmDelSrcEntry
Deleted src entry for src 192.168.52.2"

```

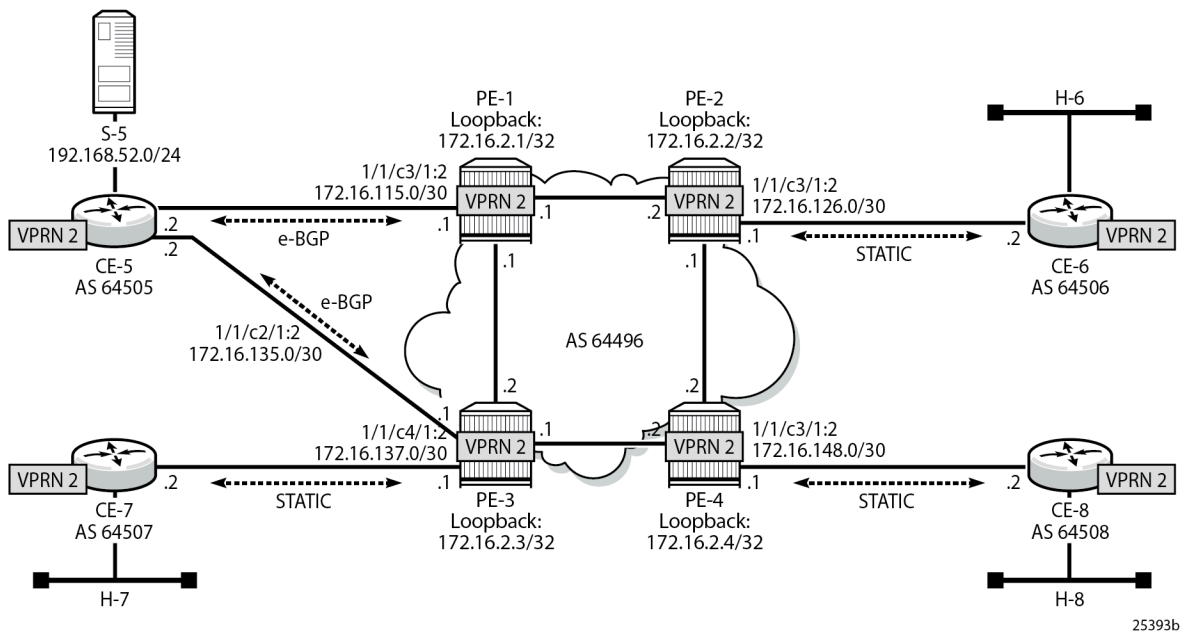
```

28 2023/10/25 09:51:48.973 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 31
  Flag: 0x90 Type: 15 Len: 27 Multiprotocol Unreachable NLRI:
    Address Family MVPN_IPV4
    Type: Source-Join Len:22 RD: 64496:201 SrcAS: 64496 Src: 192.168.52.2 Grp: 232.2.2.2
"
    
```

## MVPN Source Redundancy

So far, the multicast traffic has been streamed toward router CE-5 from a single source. For security, the source can be redundant (two sources attached to different CEs that connect to a pair of PEs). To simulate the redundancy, CE-5 has been connected to both PE-1 and PE-3, using VPRN 2, and equal cost multi-path (ECMP) is configured with the value of 2 in all PEs. With this configuration, any PE is able to reach the source through PE-1 and PE-2. The (S,G) is the same as the one used in P2MP RSVP TE (192.168.52.1, 232.2.2.2). [Figure 32: VPRN 2 Topology used for MVPN Source Redundancy](#) shows the VPRN 2 topology with the source redundancy.

Figure 32: VPRN 2 Topology used for MVPN Source Redundancy



The configuration change with respect to the previous section (P2MP RSVP-TE PMSIs) is an additional interface created in both CE-5 and PE-3 (int-CE-5-PE-3 on CE-5 and int-PE-3-CE-5 on PE-3), the addition of these interfaces to PIM and also the creation an e-BGP session between the two routers. The following is the additional configuration on PE-3 (CE-5 configuration changes are not displayed for brevity).

```

# on PE-3
configure
  service
    vprn 2
      interface "int-PE-3-CE-5" create
    
```

```

        address 172.16.35.1/30
        sap 1/1/c4/1:2 create
        exit
    exit
    bgp
        group "EXTERNAL"
            type external
            peer-as 64505
            neighbor 172.16.35.2
        exit
    exit
    no shutdown
    exit
    pim
        interface "int-PE-3-CE-5"
        exit
        no shutdown
    exit
    exit all

```

Checking the routes on PE-4, the source is reachable through PE-1 and PE-2 as ECMP is set to 2. If the configuration of the VPRN is provisioned with **auto-bind-tunnel resolution-filter rsvp resolution filter**, instead of static spoke-SDPs, the command **ignore-nh-metric** is also needed.

```
*A:PE-4# configure service vprn 2 ecmp 2
```

```
*A:PE-4# show router 2 route-table
```

```

=====
Route Table (Service: 2)
=====
Dest Prefix[Flags]                Type   Proto   Age           Pref
  Next Hop[Interface Name]                               Metric
-----
---snip---
192.168.52.0/24                   Remote BGP VPN  00h02m29s  170
    192.0.2.1 (tunneled)                               0
192.168.52.0/24                   Remote BGP VPN  00h02m29s  170
    192.0.2.3 (tunneled)                               0
-----
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
=====

```

When PE-4 receives a c-join/prune, PE-4 needs to find the **upstream multicast hop (UMH)** for the (S,G). This is the upstream multihop selection and is configurable. The values are highest-ip, hash-based, tunnel-status, and unicast-rt-pref

```

*A:PE-4# configure service vprn 2 mvpn umh-selection
- no umh-selection
- umh-selection {highest-ip|hash-based|tunnel-status|unicast-rt-pref}

```

The default is highest-ip, which is the selection of the highest /32 IP addresses (in this setup, PE-3 is preferred versus PE-1). A BGP c-join is sent with the route target equal to the VRF import extended community distributed by PE-3 for the subnet of the source (see following PE-4 debug).

```

17 2023/10/25 09:50:16.837 UTC MINOR: DEBUG #2001 vprn2 PIM[vprn 3 vprn2]
"PIM[vprn 3 vprn2]: pimSGUpStateJMribChange
SG (192.168.52.2,232.2.2.2), type <S,G> newMribNhopIp 192.0.2.3
newRpfNbrIp 192.0.2.3 newRpfType REMOTE newRpfIf 73734"

---snip---

20 2023/10/25 09:50:16.837 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.3
"Peer 1: 192.0.2.3: UPDATE
Peer 1: 192.0.2.3 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 76
  Flag: 0x90 Type: 14 Len: 33 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.4
    Type: Source-Join Len:22 RD: 64496:203 SrcAS: 64496 Src: 192.168.52.2 Grp: 232.2.2.2
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:192.0.2.3:3
"

```

The second option is hash-based, where the UMH is selected (both PEs are potentially possible UMHs) after hashing the source and group addresses of the stream. For this example, PE-3 is also preferred.

The third option, tunnel-status, is based on the status of the P2MP RSVP tunnel (not available in mLDP or PIM). The roots PE-1 and PE-3 are sending BFD messages to the leaf PE-4 (in fact this is UFD, unidirectional forwarding detection). The c-join from PE-4 for the (S,G) is sent to both PE-1 and PE-3, and in return the traffic is forwarded from both PE-1 and PE-3 for the c-group onto the I-PMSI; therefore PE-4 receives two copies of the c-(S,G) stream. By configuration, the stream from the primary PE-1 is selected by PE-4 to be forwarded to receiver H-8. If BFD messages are no longer received over the primary P2MP LSP, then the stream from the standby PE-3 is selected and forwarded to the receiver.

The configuration on PE-1 and PE-3 is similar and is as follows (only PE-3 is shown):

```

# on PE-3
configure
  service
    vprn 2
      pim shutdown
      mvpn
        auto-discovery default
        c-mcast-signaling bgp
        umh-selection tunnel-status
        provider-tunnel
          inclusive
          rsvp
            lsp-template "VRF2"
            enable-bfd-root 100
            no shutdown
          exit
        exit
      exit
    exit
  exit

```

```

        vrf-target unicast
        exit
    exit
    pim no shutdown
exit all

```

PE-1 and PE-3 are root. On PE-4, BFD is configured as leaf and the primary PE (PE-1) and backup PE (PE-3) are also provisioned:

```

# on PE-4
configure
service
  vprn 2
  pim shutdown
  mvpn
    auto-discovery default
    c-mcast-signaling bgp
    umh-selection tunnel-status
    umh-pe-backup
    umh-pe 192.0.2.1 standby 192.0.2.3
  exit
  provider-tunnel
    inclusive
    rsvp
      lsp-template "VRF2"
      enable-bfd-leaf
      no shutdown
    exit
  exit
  vrf-target unicast
  exit
exit
  pim no shutdown
exit all

```

This BFD (UFD) configuration on the root establishes a session with the leaf. The root only transmits BFD packets; it doesn't receive any.



**Note:** In the used setup, received packets and transmitted packets are not counted (value: N/A). In addition, the configured transmit interval (value: 100) is not considered; the default (value: 1000) is used instead.

```
*A:PE-1# show router 2 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
mpls-if-73736	Up	N/A	N/A
127.0.0.0	3	1000	N/A
pim	cpm-np	N/A	N/A
192.0.2.1			

```
-----
No. of BFD sessions: 1
=====
```

On PE-4, two BFD sessions are received, one from each root (BFD packets are only received):

```
*A:PE-4# show router 2 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 Multipl   Tx Intvl Rx Intvl
Protocols          Type     LAG Port  LAG ID
Loc Addr           LAG name
-----
mpls-if-73739      Up        N/A      N/A
192.0.2.3          3        N/A      1000
pim                cpm-np   N/A      N/A
192.0.2.4
mpls-if-73740      Up        N/A      N/A
192.0.2.1          3        N/A      1000
pim                cpm-np   N/A      N/A
192.0.2.4
-----
No. of BFD sessions: 2
=====
```

PE-4 delivers the multicast traffic from the primary configured UMH, PE-1. If, as an example of a failure condition, PE-1 goes down (reboot), PE-4 switches to the PE-3 P2MP LSP.

## PE VPRN Configuration and PE VPRN Multicast Configuration for Rosen MVPN using MDT AFI SAFI

In Rosen MVPN up to version 6, the default MDT is PIM sparse mode only, and there is no auto-discovery mechanism available. In SR OS Release 7.0, and later, it is possible to configure PIM SSM with auto-discovery, using AFI 1 and SAFI 5. Rosen MVPN version 7 allows use of MDT PIM SM or SSM, and auto-discovery based on AFI 1 and SAFI 66 to distribute the default MDT. Rosen MVPN version 9 adds a new MDT NLRI. SR OS has added the capability and support of MDT SAFI as specified in RFC 6037.

MDT SAFI is used to discover PEs in a specific MVPN, so that PIM SSM can be used for default MDT. The basic idea is the same as MVPN BGP auto-discovery, but it uses a different BGP SAFI. BGP messages in which AFI=1 and SAFI=66 are "MDT-SAFI" messages. The NLRI format is 8-byte-RD:IPv4-address followed by the MDT group address. The IPv4 address identifies the PE that originated this route and the RD identifies a VRF in that PE. The group address must be an IPv4 multicast group address and is used to build the P-tunnels.

All PEs attached to a specific MVPN must specify the same group-address. MDT-SAFI routes do not carry RTs and the group address is used to associate a received MDT-SAFI route with a VRF.

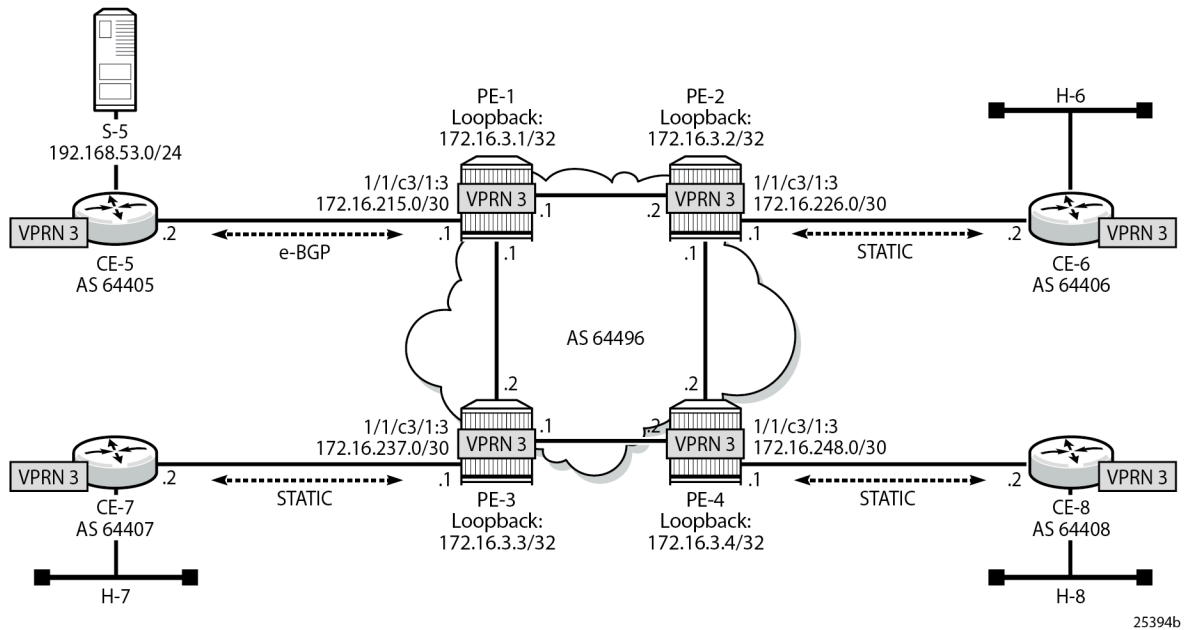
MDT SAFI can only be used when the implicit provider tunnel is PIM GRE based with a specific IPv4 group address.

For more information about the use of PIM PMSIs, see [NG-MVPN Configuration with PIM](#).



Figure 33: VPRN 3 Topology used for MDT SAFI shows the topology of VPRN 3.

Figure 33: VPRN 3 Topology used for MDT SAFI



In this scenario, there is no MPLS based PMSI, there is PIM in the core for the control plane and the data traffic is GRE encapsulated. PIM needs to be configured in the base router on interface system and on the other interfaces pointing to other PEs. PIM is used for c-signaling. In addition, auto-discovery is provisioned to use mdt-safi and a PIM SSM inclusive PMSI with address 239.1.1.1 as the default MDT. The configuration is as follows on PE-4:

```
# on PE-4
configure
router
  pim
    interface "system"
    exit
    interface "int-PE-4-PE-2"
    exit
    interface "int-PE-4-PE-3"
    exit
    no shutdown
  exit
exit
service
  vprn 3 name "3" customer 1 create
  description "PIM SSM / MDT SAFI"
  autonomous-system 64496
  interface "loopback" create
  address 172.16.3.4/32
  loopback
  exit
  interface "int-PE-4-CE-8" create
  address 172.16.248.1/30
  sap 1/1/c3/1:3 create
  exit
exit
```

```

    bgp-ipvpn
      mpls
        route-distinguisher 64496:304
        vrf-target target:64496:300
        no shutdown
      exit
    exit
  pim
    interface "loopback"
    exit
    interface "int-PE-4-CE-8"
    exit
    no shutdown
  exit
  mvpn
    auto-discovery mdt-safi
    provider-tunnel
      inclusive
      pim ssm 239.1.1.1
    exit
  exit
  vrf-target unicast
  exit
  spoke-sdp 341:3 create
  exit
  spoke-sdp 342:3 create
  exit
  spoke-sdp 343:3 create
  exit
  no shutdown
exit

```

The following debug output shows a BGP update with MDT AFI SAFI on PE-4:

```

11 2023/10/25 10:01:09.192 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 62
  Flag: 0x90 Type: 14 Len: 26 Multiprotocol Reachable NLRI:
    Address Family MDT-SAFI
    NextHop len 4 NextHop 192.0.2.4
    [MDT-SAFI] Addr 192.0.2.4, Group 239.1.1.1, RD 64496:304
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:64496:300
"

```

The following output shows the MDT-SAFI routes that have been learned at PE-4:

```

*A:PE-4# show router bgp routes mdt-safi
=====
BGP Router ID:192.0.2.4      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
Origin codes  : i - IGP, e - EGP, ? - incomplete

```

```

=====
BGP MDT-SAFI Routes
=====
Flag  Network                NextHop          Group-Addr      LocalPref  MED  IGP Cost  Label
      As-Path
-----
u*>i  64496:301:192.0.2.1    192.0.2.1      239.1.1.1      100        0    -         -
      No As-Path
u*>i  64496:302:192.0.2.2    192.0.2.2      239.1.1.1      100        0    -         -
      No As-Path
u*>i  64496:303:192.0.2.3    192.0.2.3      239.1.1.1      100        0    -         -
      No As-Path
-----
Routes : 3
=====

```

## Conclusion

This chapter provides information to configure multicast within a VPRN with next generation multicast VPN techniques. Specifically, the use of MPLS I-PMSIs (mLDP and P2MP RSVP-TE), MVPN source redundancy, and the complete set of features needed to interoperate with Rosen MVPN in live deployments are covered.

# NG-MVPN Configuration with PIM

This chapter provides information about multicast in a VPRN service.

Topics in this chapter include:

- [Applicability](#)
- [Overview](#)
- [Configuration](#)
- [Conclusion](#)

## Applicability

Initially, this chapter was written for SR OS Release 7.0.R5. The configuration in the current edition is based on SR OS Release 23.7.R2. There are no prerequisites for this configuration.

## Overview

Multicast VPN (MVPN) architectures describe a set of VRFs that support the transport of multicast traffic across a provider network.

RFC 6037 (herein referred to as Rosen MVPN) describes the use of Multicast Distribution Trees (MDTs) established between PEs within a VRF. Each VRF requires its own tree. Customer Edge (CE) routers form Protocol Independent Multicast (PIM) adjacencies with the PE, and PE-PE PIM adjacencies are formed across the multicast tree. PIM signaling and data streams are transported across the MDT. There are a number of limitations with the Rosen MVPN implementation including, but not limited to:

- Rosen MVPN requires a set of MDTs per VPN, which requires a PIM state per MDT. There is no option to aggregate MDT across multiple VPNs
- Customer signaling, PE discovery and Data MDT signaling are all PIM-based. There is no mechanism available to decouple these. Thus there is an incongruency between unicast and multicast VPNs using Rosen MVPN.
  - There is no mechanism for using MPLS to encapsulate multicast traffic in the VPN. GRE is the only encapsulation method available in Rosen MVPN.
  - Rosen MVPN multicast trees are signaled using PIM only. Next Generation MVPN (NG MVPN) allows the use of mLDP, RSVP P2MP LSPs.
  - PE to PE protocol exchanges for Rosen MVPN is achieved using PIM only. NG MVPN allows for the use of BGP signaling as per unicast Layer 3 VPNs.

NG MVPN addresses these limitations by extending the idea of the per-VRF tree, by introducing the idea of Provider Multicast Service Interfaces (PMSI). These are equivalent to the default MDTs of Rosen MVPN in that they support control plane traffic (customer multicast signaling), and the data MDTs which carry multicast data traffic streams between PEs within a multicast VRF.

NG MVPN allows the decoupling of the mechanism required to create a multicast VPN, such as PE auto-discovery (which PEs are members of which VPN), PMSI signaling (creation of tunnels between PEs) and customer multicast signaling (multicast signaling -IGMP/PIM- received from CE routers). Two types of PMSI exist:

- Inclusive (I-PMSI): contains all the PEs for an MVPN.
- Selective (S-PMSI): contains only a subset of PEs of an MVPN.

Knowledge of MPLS-VPN RFC 4364, *BGP/MPLS IP Virtual Private Networks (VPNs)*, architecture and functionality, as well as an understanding of multicast protocols, is assumed throughout.

This chapter provides configuration details required to implement the parts of NG MVPN shown in [Table 1: NG MVPN Components](#).

Table 1: NG MVPN Components

Provider Multicast Domain				Customer Multicast Domain		
I-PMSI	Auto-discovery	C-Mcast	S-PMSI Creation	PE-based RP	Anycast RP on PE	PIM SSM
PIM ASM	PIM	PIM join/leave	PIM SSM with S-PMSI join TLV	X	X	X
PIM ASM	BGP A/D	PIM join/leave	PIM SSM with S-PMSI join TLV			X

The first section of this chapter describes the common configuration required for each PE within the provider multicast domain, regardless of the MVPN PE auto-discovery or customer signaling methods. This includes IGP and VPRN service configuration.

Following the common configuration, specific MVPN configuration required for the configuration for the provider multicast domain using PIM Any Source Multicast (ASM) with auto-discovery based on PIM or BGP auto-discovery (A/D), PIM used for the customer multicast signaling and PIM Source Specific Multicast (SSM) used for the S-PMSI creation are described. The customer domain configuration covers the following three cases:

1. PIM ASM with the Rendezvous Point (RP) in the provider PE
2. PIM ASM using anycast RP on the provider RPs
3. PIM SSM

Other possible options, not covered in this section, are described in the 7450 ESS, 7750 SR, 7950 XRS, and VSR Multicast Routing Protocols Guide:

- The use of PIM SSM for the provider multicast I-PMSI.
- The use of BGP for the customer multicast signaling in the provider multicast domain.
- The provider S-PMSI creation through BGP S-PMSI A/D.
- The use of the customer RP based in the customer CE.

The use of mLDP and RSVP p2mp LSPs for the I/S-PMSI was not available in Release 7.0.

The Multicast in a VPRN II example in [NG-MVPN Configuration with MPLS](#) introduces features that were not supported in Release 7.0.R5. It provides configuration details to implement:

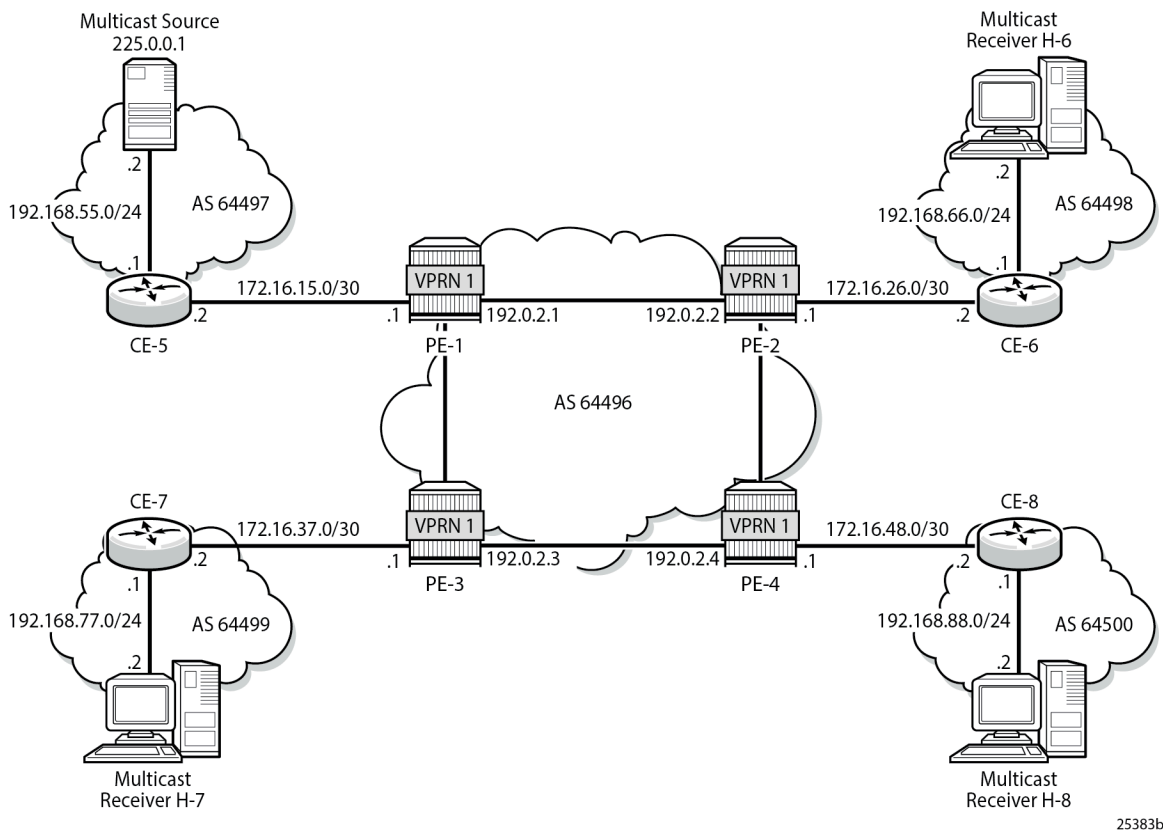
- Multicast LDP (mLDP) and RSVP-TE Point to Multi-point (P2MP) for building customer trees (C-trees) which are using MPLS instead of PIM techniques.
- MVPN source redundancy
- MDT AFI/SAFI (to fully interoperate with Cisco networks).

## References

- IETF
  - RFC 6513, *Multicast in MPLS/BGP IP VPNs*
  - RFC 6514, *BGP Encodings and Procedures for Multicast in MPLS/ BGP IP VPNs*
- 7450 ESS, 7750 SR, 7950 XRS, and VSR Layer 3 Services Guide: IES and VPRN

## Topology

Figure 34: Network Topology



The network topology is displayed in [Figure 34: Network Topology](#). The setup consists of four SR 7750s acting as Provider Edge (PE) routers within a single Autonomous System (AS).

- Full mesh IS-IS or OSPF in each AS

- LDP on all interfaces in each AS (RSVP could also be used)
- MP-iBGP sessions between the PE routers in each AS (Route Reflectors (RRs) could also be used).
- Layer 3-VPN on all PEs with identical route targets, in the form AS-number: *vprn-service-id*

Connected to each PE is a single SR OS router acting as a Customer Edge (CE) router. CE-5 has a multicast source connected, and CE-6, CE-7, and CE-8 each have a single receiver connected which receive the multicast streams from the source. In this document, each receiver is both IGMPv2 and IGMPv3 capable. If the customer domain multicast signaling plane uses Source Specific Multicasting (SSM), then an IGMPv3 receiver is configured; if Any Source Multicasting (ASM) is used, the receiver is IGMPv2 capable.

If the receiver is IGMPv3 capable, it issues IGMPv3 reports that include a list of required source addresses. The receiver joins the 232.0.0.1 multicast group.

If the receiver is only IGMPv2 capable, then it issues IGMPv2 reports which do not specify a source of the group. In this case, a Rendezvous Point is required within the PIM control plane of the multicast VRF which is source-aware. In this case, the receiver joins the 225.0.0.1 multicast group.

When the receiver wants to become a member of any group, the source address of the group must be known to the CE. As a result, the source address must be IP reachable by each CE, so it is advertised by CE-5 to the PEs with attachment circuits in VPRN1 using BGP.

Static routes are then configured on the receiver CEs to achieve IP reachability to the source address of multicast groups. In the case of PIM ASM, any RP that is configured must also be reachable from the CE.

## Multicast VPN Overview

Multicast traffic from the source is streamed toward router CE-5. Receivers connected to CE-6, CE-7 and CE-8 are interested in joining this multicast group.

All CEs are PIM enabled routers, which form a PIM adjacency with their nearest PE. The PIM adjacencies between PEs across the Provider network are achieved using I-PMSIs. I-PMSIs carry PIM control messages between PEs. Data plane traffic is transported across the I-PMSI until a configured bandwidth threshold is reached. A Selective PMSI is then signaled that carries data plane traffic. This threshold can be as low as 1kb/second and must be explicitly configured along with the S-PMSI multicast group. An S-PMSI per customer group per VPRN is configured. If no S-PMSI and threshold is configured, data traffic continues to be forwarded across the provider network within the I-PMSI.

## Configuration

The configuration is divided into the following sections:

- [Provider Common Configuration](#)
  - [PE Global Configuration](#)
  - [PE VPRN Configuration](#)
- [PE VPRN Multicast Configuration](#)
  - [Auto-Discovery within Provider Domain Using PIM](#)
  - [PIM Auto-Discovery - Customer Signaling using PIM](#)
- [PIM Auto-Discovery - Customer Signaling using PIM](#)

- [PIM Any Source Multicasting with RP at the Provider PE](#)
- [PIM Any Source Multicasting with Anycast RP at the Provider PE](#)
- [PIM Source Specific Multicasting](#)
  - [BGP Auto-Discovery - PE VPRN Multicast Configuration](#)
  - [Data Path Using Selective PMSI](#)

## Provider Common Configuration

This section describes the common configuration required for each PE within the Provider multicast domain, regardless of the MVPN PE auto-discovery or customer signaling methods. This includes IGP and VPRN service configuration.

The configuration tasks can be summarized as follows:

- [PE Global Configuration](#). This includes configuration of the Interior Gateway Protocol (IGP) (IS-IS or OSPF); configuration of link layer LDP between PEs; configuration of iBGP between PEs, to facilitate VPRN route learning; configuration of PIM.
- [PE VPRN Configuration](#). This includes configuration of basic VPRN parameters (route-distinguisher, route target communities); configuration of attachment circuits toward CEs; configuration of VRF routing protocol and any policies toward CE.
- VRF PIM and MVPN parameters - I-PMSI
- CE configuration.

## PE Global Configuration

1. On each of the PE routers, configure the appropriate router interfaces, OSPF (or IS-IS) and link layer LDP. For clarity in the following configuration steps, only the configuration for PE-1 is shown. PE-2, PE-3, and PE-4 are similar.

```
On PE-1:
configure
  router
    interface "int-PE-1-PE-2"
      address 192.168.12.1/30
      port 1/1/c1/1
    exit
    interface "int-PE-1-PE-3"
      address 192.168.13.1/30
      port 1/1/c2/1
    exit
    interface "system"
      address 192.0.2.1/32
    exit
    autonomous-system 64496
    ospf 0
      area 0.0.0.0
        interface "system"
        exit
        interface "int-PE-1-PE-2"
          interface-type point-to-point
        exit
        interface "int-PE-1-PE-3"
```



```

        interface-type point-to-point
        exit
    exit
    no shutdown
exit
ldp
    interface-parameters
        interface "int-PE-1-PE-2"
        exit
        interface "int-PE-1-PE-3"
        exit
    exit
exit
exit all

```

**2. Verify that OSPF adjacencies are formed and that LDP peer sessions are formed.**

```
*A:PE-1# show router ospf neighbor
```

```

=====
Rtr Base OSPFv2 Instance 0 Neighbors
=====
Interface-Name          Rtr Id          State    Pri  RetxQ  TTL
Area-Id
-----
int-PE-1-PE-2          192.0.2.2      Full     1    0      31
0.0.0.0
int-PE-1-PE-3          192.0.2.3      Full     1    0      33
0.0.0.0
-----
No. of Neighbors: 2
=====

```

```
*A:PE-1# show router ldp session ipv4
```

```

=====
LDP IPv4 Sessions
=====
Peer LDP Id            Adj Type  State      Msg Sent  Msg Recv  Up Time
-----
192.0.2.2:0           Link     Established  43        44        0d 00:01:29
192.0.2.3:0           Link     Established  27        29        0d 00:00:47
-----
No. of IPv4 Sessions: 2
=====

```

**3. Configure BGP between the PEs for VPRN routing.**

```

On PE-1:
configure
    router
        bgp
            group "INTERNAL"
                family vpn-ipv4
                type internal
                neighbor 192.0.2.2
                exit
                neighbor 192.0.2.3
                exit
                neighbor 192.0.2.4
                exit
            exit

```

```
no shutdown
exit all
```

- Verify that BGP sessions are established for address family VPN-IPv4.

```
*A:PE-1# 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
-----
192.0.2.2
Def. Inst          64496      5   0 00h00m33s 0/0/0 (VpnIPv4)
                   5   0
192.0.2.3
Def. Inst          64496      3   0 00h00m23s 0/0/0 (VpnIPv4)
                   3   0
192.0.2.4
Def. Inst          64496      3   0 00h00m12s 0/0/0 (VpnIPv4)
                   3   0
-----
```

- Enable PIM on all network interfaces, including the system interface. This allows the signaling of PMSIs that transport PIM signaling within each VRF.
- Each I-PMSI is signaled using PIM ASM, so a rendezvous point (RP) is required within the global PIM configuration. A static RP is used and PE-1 is selected. All PEs must be configured with this RP address.

```
On PE-1:
configure
router
pim
    interface "system"
    exit
    interface "int-PE-1-PE-2"
    exit
    interface "int-PE-1-PE-3"
    exit
    rp
        static
            address 192.0.2.1
            group-prefix 239.255.0.0/16
        exit
    exit
exit
exit all
```

- The following command shows the PIM neighbor relationships.

```
*A:PE-1# show router pim neighbor

=====
PIM Neighbor ipv4
=====
```

Interface Nbr Address	Nbr DR Prty	Up Time	Expiry Time	Hold Time
int-PE-1-PE-2 192.168.12.2	1	0d 00:00:41	0d 00:01:36	105
int-PE-1-PE-3 192.168.13.2	1	0d 00:00:29	0d 00:01:17	105
Neighbors : 2				
=====				

## PE VPRN Configuration

A VPRN (VPRN 1) is created on each PE. This is the multicast VPRN. PE-1 is the PE containing the attachment circuit toward CE-5. CE-5 is the CE nearest the source. PE-2, PE-3, and PE-4 contain attachment circuits toward CE-6, CE-7, and CE-8 respectively. CE-6 has receiving host H-6 attached; CE-7 has receiving host H-7, and CE-8 receiving host H-8.

1. Create VPRN 1 on each PE, containing a route-distinguisher and vrf-target of 64496:1. The autonomous system number is 64496. Use **auto-bind-tunnel resolution-filter ldp** for next hop tunnel route resolution.

```
On PE-1:
configure
  service
    vprn 1 name "VPRN 1" customer 1 create
      autonomous-system 64496
      bgp-ipvpn
        mpls
          auto-bind-tunnel
            resolution-filter
              ldp
            exit
          resolution filter
        exit
      route-distinguisher 64496:1
      vrf-target target:64496:1
      no shutdown
    exit
  exit
  no shutdown
exit all
```

2. Create an attachment circuit interface on PE-1 toward CE-5.

```
On PE-1:
configure
  service
    vprn 1
      interface "int-PE-1-CE-5" create
        address 172.16.15.1/30
        sap 1/1/c3/1 create
      exit
    exit
  no shutdown
exit all
```

- The source address of the multicast stream must be reachable by all routers (PEs and CEs) within the VPN. This is advertised within BGP from the CE to the PE. Create a BGP peering relationship within VPRN 1 on PE-1 with CE-5.

```
On PE-1:
configure
  service
    vprn 1
      bgp
        group "EXTERNAL"
          type external
          peer-as 64497
          neighbor 172.16.15.2
        exit
      exit
    exit
  no shutdown
exit all
```

- On CE-5, create a VPRN to support the connection of the source to the CE and to connect the CE to the PE. Two attachment circuits are required, as well as a BGP peering relationship with the PE. This uses a default address family of **ipv4**.

(A pair of IES services could also be used to provide the attachment circuits.)

```
On CE-5:
configure
  service
    vprn 1 name "VPRN 1" customer 1 create
      autonomous-system 64497
      bgp-ipvpn
        mpls
          route-distinguisher 64497:1
          no shutdown
        exit
      exit
    interface "int-CE-5-PE-1" create
      address 172.16.15.2/30
      sap 1/1/c1/1 create
    exit
    interface "int-CE-5-S-5" create
      address 192.168.55.1/24
      sap 1/1/c3/1 create
    exit
  bgp
    group "EXTERNAL"
      type external
      peer-as 64496
      neighbor 172.16.15.1
    exit
  exit
  no shutdown
exit all
```

- The following BGP summaries show that the PE-CE BGP peer relationship between CE-5 and PE-1 is established for address family IPv4:

```
*A:CE-5# show router 1 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
-----
172.16.15.1
1                  64496      5   0 00h00m53s 0/0/0 (IPv4)
                  5         0
-----
```

```
*A:PE-1# show router 1 bgp summary
=====
BGP Router ID:192.0.2.1      AS:64496      Local AS:64496
=====
BGP Admin State      : Up          BGP Oper State      : Up
Total Peer Groups    : 1            Total Peers         : 1
Current Internal Groups : 1          Max Internal Groups : 1
Total BGP Paths       : 9            Total Path Memory   : 3168

Total IPv4 Remote Rts : 0            Total IPv4 Rem. Active Rts : 0
---snip---

=====
BGP Summary
=====
Legend : D - Dynamic Neighbor
=====
Neighbor
Description
ServiceId          AS PktRcvd InQ Up/Down  State|Rcv/Act/Sent (Addr Family)
                  PktSent OutQ
-----
172.16.15.2
                  64497      5   0 00h00m50s 0/0/0 (IPv4)
                  5         0
-----
```

- For the CE connecting to the source to be advertised within BGP, a route policy is required. The subnet containing the multicast source is 192.168.55.0/24, so a prefix-list can be used to define a match, and then used within a route policy to inject into BGP.

```
On CE-5:
configure
router
policy-options
begin
prefix-list "SOURCE-PREFIX"
prefix 192.168.55.0/24 exact
exit
policy-statement "EXPORT-SOURCE-PREFIX-T0-BGP"
entry 10
from
prefix-list "SOURCE-PREFIX"
exit
to
protocol bgp
```

```

                                exit
                                action accept
                                exit
                                exit
                                exit
                                commit
                                exit all
    
```

7. Apply this policy as an export policy within the **bgp** context.

```

On CE-5:
configure
  service
    vprn 1
      bgp
        export "EXPORT-SOURCE-PREFIX-T0-BGP"
      exit
    no shutdown
  exit all
    
```

This results in the 192.168.55.0/24 subnet being seen in the BGP RIB OUT Entries on CE-5.

```

*A:CE-5# show router 1 bgp routes 192.168.55.0/24 hunt
=====
BGP Router ID:192.0.2.5      AS:64497      Local AS:64497
=====
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 IPv4 Routes
=====
RIB In Entries
-----snip-----
RIB Out Entries
-----
Network       : 192.168.55.0/24
NextHop       : 172.16.15.2
Path Id       : None
To            : 172.16.15.1
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 : 192.0.2.1
Origin        : IGP
AS-Path       : 64497
Route Tag     : 0
Neighbor-AS   : 64497
DB Orig Val   : N/A                    Final Orig Val : N/A
Source Class  : 0                       Dest Class     : 0
=====
    
```

```
Routes : 2
```

It is also seen in the PE-1 VRF 1 FIB:

```
*A:PE-1# show router 1 route-table
```

```
=====
```

```
Route Table (Service: 1)
```

```
=====
```

Dest Prefix[Flags] Next Hop[Interface Name]	Type	Proto	Age Metric	Pref
172.16.15.0/30 int-PE-1-CE-5	Local	Local	00h04m48s 0	0
172.16.26.0/30 192.0.2.2 (tunneled)	Remote	BGP VPN	00h04m11s 1	170
172.16.37.0/30 192.0.2.3 (tunneled)	Remote	BGP VPN	00h04m18s 1	170
172.16.48.0/30 192.0.2.4 (tunneled)	Remote	BGP VPN	00h04m16s 2	170
<b>192.168.55.0/24</b> 172.16.15.2	<b>Remote</b>	<b>BGP</b>	00h01m08s 0	170

```
-----
```

```
No. of Routes: 5
```

```
Flags: n = Number of times nexthop is repeated
```

```
      B = BGP backup route available
```

```
      L = LFA nexthop available
```

```
      S = Sticky ECMP requested
```

```
=====
```

This prefix is also automatically advertised within the BGP VPRN to all other PEs, and is installed in VRF 1. For example, on PE-2:

```
*A:PE-2# show router 1 route-table
```

```
=====
```

```
Route Table (Service: 1)
```

```
=====
```

Dest Prefix[Flags] Next Hop[Interface Name]	Type	Proto	Age Metric	Pref
172.16.15.0/30 192.0.2.1 (tunneled)	Remote	BGP VPN	00h04m37s 1	170
172.16.26.0/30 int-PE-2-CE-6	Local	Local	00h04m41s 0	0
172.16.37.0/30 192.0.2.3 (tunneled)	Remote	BGP VPN	00h04m05s 2	170
172.16.48.0/30 192.0.2.4 (tunneled)	Remote	BGP VPN	00h04m12s 1	170
<b>192.168.55.0/24</b> 192.0.2.1 (tunneled)	<b>Remote</b>	<b>BGP VPN</b>	00h01m03s 1	170

```
-----
```

```
No. of Routes: 5
```

```
Flags: n = Number of times nexthop is repeated
```

```
      B = BGP backup route available
```

```
      L = LFA nexthop available
```

```
      S = Sticky ECMP requested
```

```
=====
```

Each CE containing the multicast receivers must be able to reach the source. The following shows the VPRN configuration of CE-6 containing an interface toward PE-2 and an interface toward receiving host H-6. A static route suffices and is configured with next hop of the PE-2 PE-CE interface.

```
On CE-6:
configure
service
  vprn 1 name "VPRN 1" customer 1 create
  bgp-ipvpn
  mpls
    route-distinguisher 64498:1
    no shutdown
  exit
exit
interface "int-CE-6-H-6" create
  address 192.168.66.1/24
  sap 1/1/c2/1 create
  exit
exit
interface "int-CE-6-PE-2" create
  address 172.16.26.2/30
  sap 1/1/c1/1 create
  exit
exit
static-route-entry 192.168.55.0/24
  next-hop 172.16.26.1
  no shutdown
  exit
exit
no shutdown
exit all
```

## PE VPRN Multicast Configuration

This section gives details of the VPRN configuration that allows the support of multicasting.

Sub-sections include:

1. Auto-discovery - This is the mechanism by which each PE advertises the presence of an MVPN to other PEs. This can be achieved using PIM or using BGP. This section covers PIM auto-discovery (auto-discovery using BGP is shown later).
2. Customer domain signaling - This discusses the mechanism of transporting customer signaling.
3. Data plane connectivity - This is the signaling of S-PMSIs within the provider domain to carry each individual customer multicast stream.

This chapter describes the PIM and BGP auto-discovery mechanisms in detail. For each of these, there is an example of customer domain signaling. For completion, a single example of S-PMSI creation is also shown.



## Auto-Discovery within Provider Domain Using PIM

Each PE advertises its membership of a multicast VPN using PIM through the configuration of an Inclusive PMSI (I-PMSI). This is a multicast group that is common to each VPRN. The configuration is identical for all PEs and is as follows:

```
On all PEs:
configure
  service
    vprn 1
      mvpn
        provider-tunnel
          inclusive
            pim asm 239.255.255.1
          exit
        exit
      exit
    exit
  no shutdown
exit all
```

The multicast group address used for the PMSI must be the same on all PEs for this VPRN instance.

Verify that PIM in the Global Routing Table (GRT) has signaled the I-PMSIs.

For the PE acting as the RP for global PIM:

```
*A:PE-1# show router pim group

=====
Legend:  A = Active   S = Standby
=====
PIM Groups ipv4
=====
Group Address          Type          Spt Bit   Inc Intf   No.0ifs
Source Address         RP            State     Inc Intf(S)
-----
239.255.255.1          (*,G)                3
*
192.0.2.1              (S,G)            spt      system    3
192.0.2.1              192.0.2.1
239.255.255.1          (S,G)            spt      int-PE-1-PE-2  3
192.0.2.2              192.0.2.1
239.255.255.1          (S,G)            spt      int-PE-1-PE-3  3
192.0.2.3              192.0.2.1
239.255.255.1          (S,G)            spt      int-PE-1-PE-2  2
192.0.2.4              192.0.2.1
-----
Groups : 5
=====
```

This shows an incoming (S,G) join from all other PEs within the multicast VRF, plus an outgoing (\*,G) join to the same PEs.

PE-3 has the following PIM groups:

```
*A:PE-3# show router pim group

=====
Legend:  A = Active   S = Standby
=====
```

```

=====
PIM Groups ipv4
=====
Group Address          Type          Spt Bit   Inc Intf   No.Oifs
Source Address        RP           State     Inc Intf(S)
-----
239.255.255.1        (*,G)                int-PE-3-PE-1  1
*                    192.0.2.1
239.255.255.1        (S,G)          spt       system      2
192.0.2.3            192.0.2.1
-----
Groups : 2
=====

```

This shows an (S,G) join toward the RP at 192.0.2.1, plus a (\*,G) join from the RP. These represent the outgoing and incoming PIM interfaces for the VRF.

This results in a series of PIM neighbors through the I-PMSIs within the VRF, which are maintained using PIM hellos.

```

*A:PE-1# show router 1 pim neighbor
=====
PIM Neighbor ipv4
=====
Interface          Nbr DR Prty   Up Time      Expiry Time   Hold Time
Nbr Address
-----
int-PE-1-CE-5      1              0d 00:00:45  0d 00:01:31  105
172.16.15.2
1-mt-239.255.255.1 1              0d 00:01:20  0d 00:01:28  105
192.0.2.2
1-mt-239.255.255.1 1              0d 00:01:07  0d 00:01:41  105
192.0.2.3
1-mt-239.255.255.1 1              0d 00:00:56  0d 00:01:22  105
192.0.2.4
-----
Neighbors : 4
=====

```

## PIM Auto-Discovery - Customer Signaling using PIM

Consider now how the signaling plane of the customer domain is dealt with at the provider domain.

The customer domain configuration covers the following three cases:

1. [PIM Any Source Multicasting with RP at the Provider PE](#)
2. [PIM Any Source Multicasting with Anycast RP at the Provider PE](#)
3. [PIM Source Specific Multicasting](#)

## PIM Any Source Multicasting with RP at the Provider PE

Each PE connects to a CE which is part of the multicast VRF, so it is necessary to enable PIM on each interface containing an attachment circuit toward a CE, and to configure the I-PMSI multicast tunnel for the VRF.

There is a requirement for an RP, because customer multicast signaling is PIM-ASM.

The RP for the customer multicast is on PE-2. To facilitate this, a loopback interface (called "RP") is created within the **vprn 1** context of PE-2, and is advertised to all PEs. It must also be a PIM enabled interface.

The additional configuration for the RP on PE-2 is the following:

```
On PE-2:
configure
  service
    vprn 1
      interface "RP" create
        address 10.2.3.5/32
        loopback
      exit
    pim
      interface "RP"
        exit
        no shutdown
        rp
          static
            address 10.2.3.5
            group-prefix 225.0.0.0/8
          exit
        exit
      exit
    exit
  no shutdown
exit all
```

The RP must also be configured on each of the PEs and CEs.

On PE-3, the PIM configuration in VPRN 1 is as follows:

```
On PE-3:
configure
  service
    vprn 1
      pim
        interface "int-PE-3-CE-7"
          exit
          no shutdown
          rp
            static
              address 10.2.3.5
              group-prefix 225.0.0.0/8
            exit
          exit
        exit
      exit
    no shutdown
exit all
```

The configuration on the other nodes is similar; only the interfaces are different.

## Customer Edge Router Multicast Configuration

Each CE router has a PIM neighbor peer relationship with its nearest PE.

The CE router (CE-5) containing the source has PIM enabled on the interface connected to the source. It also has a static RP entry, as the incoming sources need to be registered with the RP.

```
On CE-5:
configure
  service
    vprn 1
      pim
        interface "int-CE-5-PE-1"
        exit
        interface "int-CE-5-S-5"
        exit
        no shutdown
        rp
          static
            address 10.2.3.5
            group-prefix 225.0.0.0/8
          exit
        exit
      exit
    exit
  no shutdown
exit all
```

The CE containing the receivers has IGMP enabled on the interface connected to the receivers. Again, there needs to be an RP configured, because the router needs to issue PIM joins to the RP. The additional configuration in VPRN 1 on CE-6 is as follows:

```
On CE-6:
configure
  service
    vprn 1
      static-route-entry 10.0.0.0/8
        next-hop 172.16.26.1
        no shutdown
      exit
    exit
    static-route-entry 192.168.55.0/24
      next-hop 172.16.26.1
      no shutdown
    exit
    igmp
      interface "int-CE-6-H-6"
      exit
    exit
    pim
      interface "int-CE-6-PE-2"
      exit
      no shutdown
      rp
        static
          address 10.2.3.5
          group-prefix 225.0.0.0/8
        exit
      exit
    exit
  no shutdown
exit all
```

## Traffic Flow

The source sends a multicast stream using group address 225.0.0.1 toward CE-5. As the group matches the group address in the static RP configuration, the router sends a register join toward the RP. At this time, no receivers are interested in the group, so there are no entries in the Outgoing Interface List (OIL), and the number of outgoing interfaces (OIFs) is zero.

The PIM status of CE-5 within VPRN 1 is as follows:

```
*A:CE-5# show router 1 pim group
=====
Legend:  A = Active   S = Standby
=====
PIM Groups ipv4
=====
Group Address          Type          Spt Bit  Inc Intf      No.Oifs
Source Address        RP
-----
225.0.0.1              (S,G)                State      Inc Intf(S)
192.168.55.2          10.2.3.5          int-CE-5-S-5  0
-----
Groups : 1
=====
```

The receiver H-6 connected to CE-6, wants to join the group 225.0.0.1, and sends an IGMPv2 report toward CE-6. CE-6 recognizes the report, which contains no source.

```
*A:CE-6# show router 1 igmp group
=====
IGMP Interface Groups
=====
(*,225.0.0.1)          UpTime: 0d 00:00:15
  Fwd List : int-CE-6-H-6
-----
Entries : 1
=====
IGMP Host Groups
=====
No Matching Entries
=====
IGMP SAP Groups
=====
No Matching Entries
=====
IGMP SLA Profile Instance Groups
=====
No Matching Entries
=====
```

CE-6 is not aware of the source of the group so initiates a (\*,G) PIM join toward the RP.

At the RP, the following (\*,G) join is received:

```
*A:PE-2# show router 1 pim group 225.0.0.1 type starg detail
=====
PIM Source Group ipv4
=====
```

```

Group Address      : 225.0.0.1
Source Address    : *
RP Address        : 10.2.3.5
Advt Router      : 192.0.2.2
Flags             :
Mode             : sparse
MRIB Next Hop    :
MRIB Src Flags   : self
Keepalive Timer  : Not Running
Up Time          : 0d 00:00:27      Resolved By      : rtable-u

Up JP State       : Joined          Up JP Expiry     : 0d 00:00:33
Up JP Rpt        : Not Joined StarG Up JP Rpt Override : 0d 00:00:00

Rpf Neighbor     :
Incoming Intf    :
Outgoing Intf List : int-PE-2-CE-6

Curr Fwding Rate : 0.000 kbps
Forwarded Packets : 0              Discarded Packets : 0
Forwarded Octets  : 0              RPF Mismatches    : 0
Spt threshold     : 0 kbps          ECMP opt threshold : 7
Admin bandwidth   : 1 kbps
-----
Groups : 1
=====

```

The RP can now forward traffic from itself toward CE-6, as the outgoing interface is seen as int-PE-2-CE-6.

CE-6 is now able to determine the source from the traffic stream, so it initiates a Reverse Path Forwarding (RPF) lookup of the source address in the route table, and issues an (S,G) PIM join toward the source.

The join is propagated across the provider network, from PE-2 toward PE-1 which is the resolved RPF next hop for the source.

```

*A:PE-1# show router 1 pim group detail

=====
PIM Source Group ipv4
=====
Group Address      : 225.0.0.1
Source Address     : 192.168.55.2
RP Address        : 10.2.3.5
Advt Router      : 172.16.15.2
Flags             : spt              Type             : (S,G)
Mode             : sparse
MRIB Next Hop    : 172.16.15.2
MRIB Src Flags   : remote
Keepalive Timer Exp: 0d 00:02:06
Up Time          : 0d 00:01:23      Resolved By      : rtable-u

Up JP State       : Joined          Up JP Expiry     : 0d 00:00:36
Up JP Rpt        : Not Joined StarG Up JP Rpt Override : 0d 00:00:00

Register State    : No Info
Reg From Anycast RP: No

Rpf Neighbor     : 172.16.15.2
Incoming Intf    : int-PE-1-CE-5
Outgoing Intf List : 1-mt-239.255.255.1

Curr Fwding Rate : 4811.800 kbps
Forwarded Packets : 51324          Discarded Packets : 0

```

```
Forwarded Octets : 50400168      RPF Mismatches : 0
Spt threshold   : 0 kbps        ECMP opt threshold : 7
Admin bandwidth : 1 kbps
-----
```

```
Groups : 1
=====
```

The outgoing interface is the I-PMSI: 1-mt-239.255.255.1.

The join is received by CE-5, which contains the subnet of the source.

CE-5 now recognizes the multicast group as a valid stream. This becomes the root of the shortest path tree for the group.

```
*A:CE-5# show router 1 pim group
```

```
Legend:  A = Active  S = Standby
=====
```

```
PIM Groups ipv4
=====
```

Group Address Source Address	Type RP	Spt Bit State	Inc Intf Inc Intf(S)	No.0ifs
225.0.0.1 192.168.55.2	(S,G) 10.2.3.5	spt	int-CE-5-S-5	1

```
Groups : 1
=====
```

For completion, consider a second receiver H-7 interested in group 225.0.0.1. The IGMPv2 report is translated into a (\*,G) PIM join at CE-7 toward the RP.

```
*A:CE-7# show router 1 pim group type starg
```

```
Legend:  A = Active  S = Standby
=====
```

```
PIM Groups ipv4
=====
```

Group Address Source Address	Type RP	Spt Bit State	Inc Intf Inc Intf(S)	No.0ifs
225.0.0.1 *	(* ,G) 10.2.3.5		int-CE-7-PE-3	1

```
Groups : 1
=====
```

At the RP (PE-2), there is now a second interface in the OIL.

```
*A:PE-2# show router 1 pim group 225.0.0.1 type starg detail
```

```
PIM Source Group ipv4
=====
```

```
Group Address : 225.0.0.1
Source Address : *
RP Address    : 10.2.3.5
Advt Router   : 192.0.2.2
Flags        :
Mode         : sparse
Type         : (* ,G)
```

```

MRIB Next Hop      :
MRIB Src Flags     : self
Keepalive Timer    : Not Running
Up Time           : 0d 00:02:20      Resolved By       : rtable-u

Up JP State        : Joined           Up JP Expiry      : 0d 00:00:40
Up JP Rpt         : Not Joined StarG Up JP Rpt Override : 0d 00:00:00

Rpf Neighbor      :
Incoming Intf     :
Outgoing Intf List : int-PE-2-CE-6, 1-mt-239.255.255.1

Curr Fwding Rate  : 0.000 kbps
Forwarded Packets : 0                Discarded Packets : 0
Forwarded Octets  : 0                RPF Mismatches    : 0
Spt threshold     : 0 kbps           ECMP opt threshold : 7
Admin bandwidth   : 1 kbps
-----
Groups : 1
=====

```

The second interface is the I-PMSI, which is the multicast tunnel toward all other PEs. At PE-3, the (\*,G) join has the I-PMSI as an incoming interface, and the PE-CE interface as the outgoing interface.

```

*A:PE-3# show router 1 pim group type starg detail
=====
PIM Source Group ipv4
=====
Group Address      : 225.0.0.1
Source Address     : *
RP Address         : 10.2.3.5
Advt Router       : 192.0.2.2
Flags              :
Type               : (*,G)
Mode               : sparse
MRIB Next Hop     : 192.0.2.2
MRIB Src Flags    : remote
Keepalive Timer   : Not Running
Up Time           : 0d 00:00:46      Resolved By       : rtable-u

Up JP State        : Joined           Up JP Expiry      : 0d 00:00:14
Up JP Rpt         : Not Joined StarG Up JP Rpt Override : 0d 00:00:00

Rpf Neighbor      : 192.0.2.2
Incoming Intf   : 1-mt-239.255.255.1
Outgoing Intf List : int-PE-3-CE-7

Curr Fwding Rate  : 0.000 kbps
Forwarded Packets : 6                Discarded Packets : 0
Forwarded Octets  : 5892             RPF Mismatches    : 0
Spt threshold     : 0 kbps           ECMP opt threshold : 7
Admin bandwidth   : 1 kbps
-----
Groups : 1
=====

```

Again, when the CE receives traffic from the group, it can use the source address in the packet to initiate an (S,G) join toward the source to join the Shortest Path Tree (SPT).

```

*A:CE-7# show router 1 pim group type sg detail
=====

```



```

PIM Source Group ipv4
=====
Group Address      : 225.0.0.1
Source Address     : 192.168.55.2
RP Address         : 10.2.3.5
Advt Router       :
Flags             : spt                Type           : (S,G)
Mode              : sparse
MRIB Next Hop     : 172.16.37.1
MRIB Src Flags    : remote
Keepalive Timer Exp: 0d 00:01:50
Up Time           : 0d 00:01:42      Resolved By      : rtable-u

Up JP State       : Joined           Up JP Expiry     : 0d 00:00:18
Up JP Rpt        : Not Pruned       Up JP Rpt Override : 0d 00:00:00

Register State    : No Info
Reg From Anycast RP: No

Rpf Neighbor      : 172.16.37.1
Incoming Intf     : int-CE-7-PE-3
Outgoing Intf List : int-CE-7-H-7

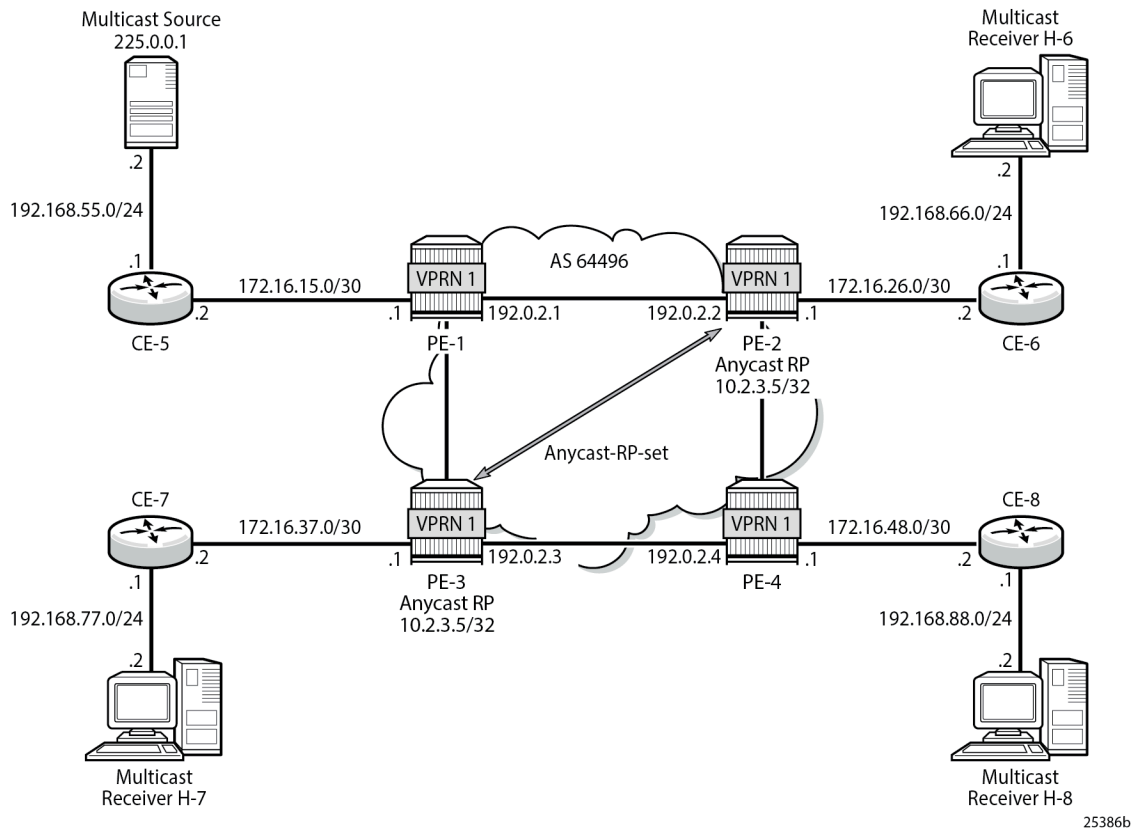
Curr Fwding Rate  : 4815.728 kbps
Forwarded Packets : 62782             Discarded Packets : 0
Forwarded Octets  : 61651924        RPF Mismatches    : 0
Spt threshold     : 0 kbps           ECMP opt threshold : 7
Admin bandwidth   : 1 kbps
-----
Groups : 1
=====

```

## PIM Any Source Multicasting with Anycast RP at the Provider PE

The example topology for anycast RP is shown in [Figure 35: Example Topology for Anycast RP](#). The setup consists of four SR OS routers acting as Provider Edge (PE) routers within a single Autonomous System (AS).

Figure 35: Example Topology for Anycast RP



Connected to each PE is a single SR OS router acting as a Customer Edge (CE) router. CE-5 has a single multicast source connected, and PEs 2 to 4 each have a single receiver connected which receives the multicast stream from the source. In this section, each receiver is IGMPv2 capable, and issues IGMPv2 reports. An RP is required by the C-signaling plane to resolve each (\*,G) group state into an (S,G) state. In this case, two RPs are chosen to form an Anycast set to resolve each (\*,G) group into an (S,G) state.

Multicast traffic from the source group 225.0.0.1 is streamed toward router CE-5. Receivers connected to PE-2, PE-3 and PE-4 are interested in joining this multicast group.

### Anycast RP - PE VPRN Configuration

As previously stated, there is a requirement for an RP, as customer multicast signaling is PIM-ASM and IGMPv2.

In this case, an anycast RP is used. This is configured on PE-2 and PE-3, and an anycast set is created.

As each PE contains a CE which is part of the multicast VRF, it is necessary to enable PIM on each interface containing an attachment circuit toward a CE, and to configure the I-PMSI multicast tunnel for the VRF.

The following shows the VPRN configuration for PE-2 containing the RP and anycast RP configuration. The loopback interface "lo1" is used for inter-RP communication:

```
On PE-2:
configure
  service
    vprn 1
      interface "RP" create
        address 10.2.3.5/32
        loopback
      exit
      interface "lo1" create
        address 10.0.0.2/32
        loopback
      exit
    pim
      interface "int-PE-2-CE-6"
      exit
      interface "RP"
      exit
      interface "lo1"
      exit
      no shutdown
      rp
        static
          address 10.2.3.5
          group-prefix 225.0.0.0/8
        exit
      exit
      anycast 10.2.3.5
        rp-set-peer 10.0.0.2
        rp-set-peer 10.0.0.3
      exit
    exit
  exit
  no shutdown
exit all
```

Similarly, the VPRN configuration for PE-3 is:

```
On PE-3:
configure
  service
    vprn 1
      interface "RP" create
        address 10.2.3.5/32
        loopback
      exit
      interface "lo1" create
        address 10.0.0.3/32
        loopback
      exit
    pim
      interface "int-PE-3-CE-7"
      exit
      interface "RP"
      exit
      interface "lo1"
      exit
      no shutdown
      rp
        static
          address 10.2.3.5
```

```

        group-prefix 225.0.0.0/8
        exit
    exit
    anycast 10.2.3.5
        rp-set-peer 10.0.0.2
        rp-set-peer 10.0.0.3
    exit
exit
exit
no shutdown
exit all

```

As previously stated, there is a requirement for an RP, as customer multicast signaling is PIM-ASM and IGMPv2.

In this case, an anycast RP is used. This is configured on PE-2 and PE-3, and an anycast set is created. The anycast address is 10.2.3.5/32 and is created as an interface called "RP" on both PE-2 and PE-3.

An additional loopback interface, called "lo1" is created on each VPRN on PEs containing the anycast address. These are used as source addresses for communication between the routers within the RP set. These addresses are automatically advertised to all PEs as VPN-IPv4 addresses, and are installed in the VRF 1 forwarding table of all PEs containing VPRN 1.

Note: All routers containing RP must have their own loopback address included in the RP set as well as all peer routers.

The multicast group address used for the Inclusive PMSI is chosen to be 239.255.255.1 and must be the same on all PEs for this VPRN instance. This is analogous to the MDT within the Rosen MVPN implementation.

```

On PE-2:
configure
  service
    vprn 1
      mvpn
        provider-tunnel
          inclusive
            pim asm 239.255.255.1
          exit
        exit
      exit
    exit
  no shutdown
exit all

```

Verify that PIM in the global routing table (GRT) has signaled the I-PMSIs.

For the PE acting as the RP for global PIM:

```

A:PE-1# show router pim group
=====
Legend:  A = Active   S = Standby
=====
PIM Groups ipv4
=====
Group Address          Type          Spt Bit  Inc Intf  No.0ifs
Source Address         RP           State    Inc Intf(S)
-----
239.255.255.1         (*,G)                3
*                      192.0.2.1

```

```

239.255.255.1      (S,G)          spt    system      3
192.0.2.1         192.0.2.1
239.255.255.1      (S,G)          spt    int-PE-1-PE-2  3
192.0.2.2         192.0.2.1
239.255.255.1      (S,G)          spt    int-PE-1-PE-3  3
192.0.2.3         192.0.2.1
239.255.255.1      (S,G)          spt    int-PE-1-PE-2  2
192.0.2.4         192.0.2.1
-----
Groups : 5
=====

```

PE-3 has:

```

*A:PE-3# show router pim group
=====
Legend:  A = Active  S = Standby
=====
PIM Groups ipv4
=====
Group Address          Type          Spt Bit   Inc Intf      No.0ifs
Source Address         RP           State     Inc Intf(S)
-----
239.255.255.1         (*,G)                int-PE-3-PE-1  1
*                      192.0.2.1
239.255.255.1         (S,G)          spt       system        2
192.0.2.3             192.0.2.1
-----
Groups : 2
=====

```

This shows a (S,G) join toward the RP at 192.0.2.1, plus a (\*,G) join from RP. These represent the outgoing and incoming PIM interfaces for the VRF.

This results in a series of PIM neighbors through the I-PMSIs within the VRF, which are maintained using PIM hellos.

```

A:PE-1# show router 1 pim neighbor
=====
PIM Neighbor ipv4
=====
Interface          Nbr DR Prty   Up Time      Expiry Time  Hold Time
Nbr Address
-----
int-PE-1-CE-5      1              0d 00:14:59  0d 00:01:17  105
172.16.15.2
1-mt-239.255.255.1 1              0d 00:15:35  0d 00:01:44  105
192.0.2.2
1-mt-239.255.255.1 1              0d 00:15:22  0d 00:01:27  105
192.0.2.3
1-mt-239.255.255.1 1              0d 00:15:10  0d 00:01:38  105
192.0.2.4
-----
Neighbors : 4
=====

```

Verify PIM RP set on PE-2 (similar for PE-3):

```

*A:PE-2# show router 1 pim anycast

```

```

=====
PIM Anycast RP Entries ipv4
=====
Anycast RP                      Anycast RP Peer
-----
10.2.3.5                        10.0.0.2
                                10.0.0.3
-----
PIM Anycast RP Entries : 2
=====

```

## Anycast RP - Customer Edge Router Multicast Configuration

Each CE router has a PIM neighbor peer relationship with its nearest PE.

The CE router (CE-5) containing the source has PIM enabled on the interface connected to the source.

```

On CE-5:
configure
  service
    vprn 1
      pim
        interface "int-CE-5-PE-1"
          exit
        interface "int-CE-5-S-5"
          exit
          no shutdown
          rp
            static
              address 10.2.3.5
              group-prefix 225.0.0.0/8
            exit
          exit
        exit
      exit
    no shutdown
  exit all

```

The CE containing the receivers has IGMP enabled on the interface connected to the receivers.

```

On CE-6:
configure
  service
    vprn 1
      igmp
        interface "int-CE-6-H-6"
          exit
        exit
      no shutdown
    exit all

```

## Traffic Flow

Figure 36: IGMP and PIM Control Messaging Schematic

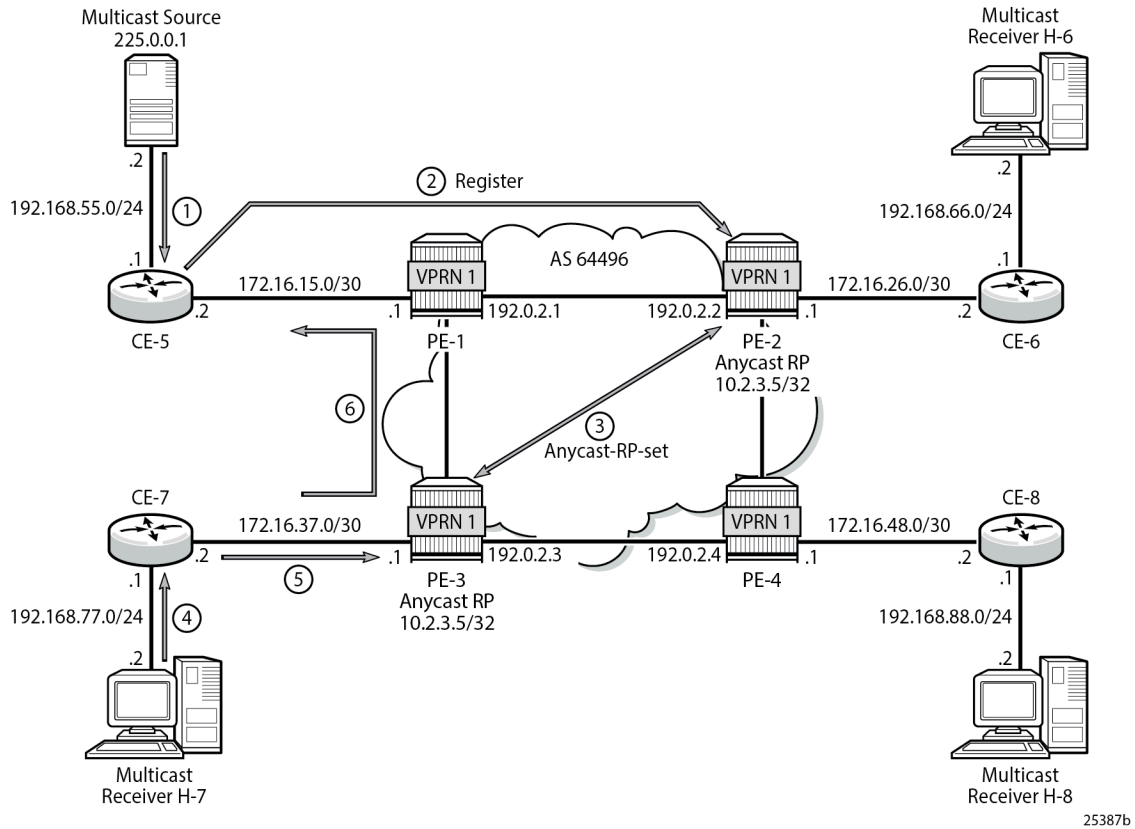


Figure 36: IGMP and PIM Control Messaging Schematic shows the sequence of IGMP and PIM control messaging.

1. The source multicasts a stream with group address 225.0.0.1 toward CE-5.
2. CE-5 matches the group with the group address prefix in the static RP configuration and sends a register message toward the RP.

```
A:CE-5# show router 1 pim group detail
```

```
=====
PIM Source Group ipv4
=====
```

```
Group Address      : 225.0.0.1
Source Address     : 192.168.55.2
RP Address         : 10.2.3.5
Advt Router       : 192.0.2.5
Flags              :
Mode               : sparse
MRIB Next Hop     : 192.168.55.2
MRIB Src Flags    : direct
Keepalive Timer Exp: 0d 00:03:05
Up Time           : 0d 00:12:37
Type               : (S,G)
Resolved By       : rtable-u
```

```

Up JP State      : Not Joined      Up JP Expiry      : 0d 00:00:00
Up JP Rpt       : Not Joined StarG Up JP Rpt Override : 0d 00:00:00

Register State   : Pruned           Register Stop Exp : 0d 00:00:40
Reg From Anycast RP: No

Rpf Neighbor    : 192.168.55.2
Incoming Intf   : int-CE-5-S-5
Outgoing Intf List :

Curr Fwding Rate : 4815.728 kbps
Forwarded Packets : 364640
Forwarded Octets  : 358076480
Spt threshold    : 0 kbps
Admin bandwidth  : 1 kbps
Discarded Packets : 0
RPF Mismatches   : 0
ECMP opt threshold : 7
-----
Groups : 1
=====
    
```

The register message is sent to the nearest RP, the RP with the lowest IGP cost.

When the register is sent through PE-1, it is PE-1 that determines which RP receives the message.

```

A:PE-1# show router 1 route-table 10.2.3.5/32
=====
Route Table (Service: 1)
=====
Dest Prefix[Flags]          Type   Proto   Age           Pref
  Next Hop[Interface Name]                               Metric
-----
10.2.3.5/32                 Remote BGP VPN 00h02m45s 170
  192.0.2.2 (tunneled)                               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
=====
    
```

The PE which receives the register is 192.0.2.2 (PE-2). The PIM group status on PE-2 is:

```

*A:PE-2# show router 1 pim group
=====
Legend:  A = Active   S = Standby
=====
PIM Groups ipv4
=====
Group Address          Type           Spt Bit   Inc Intf      No.Oifs
  Source Address       RP              State      Inc Intf(S)
-----
225.0.0.1              (S,G)          1-mt-239.255.* 0
  192.168.55.2         10.2.3.5
-----
Groups : 1
=====
* indicates that the corresponding row element may have been truncated.
    
```



This shows that RP is aware of the (S,G) status of the group 225.0.0.1, and becomes a root of a shared tree for this group. The Outgoing Interface List (OIL) is empty.

- PE-2 now sends a register message to all other RPs within the anycast set, in this case to PE-3 (which has VPRN 1 containing address 10.0.0.3).

The PIM status of the group 225.0.0.1 on PE-3 is:

```
*A:PE-3# show router 1 pim group
=====
Legend:  A = Active   S = Standby
=====
PIM Groups ipv4
=====
Group Address          Type          Spt Bit  Inc Intf      No.0ifs
Source Address         RP           State    Inc Intf(S)
-----
225.0.0.1              (S,G)                1-mt-239.255.* 0
192.168.55.2          10.2.3.5
-----
Groups : 1
=====
* indicates that the corresponding row element may have been truncated.
```

Now both PEs within the RP set for VPRN have an (S,G) state for 225.0.0.1.

- The receiver H-7 wants to join the group 225.0.0.1, and sends in an IGMPv2 report toward CE-7. CE-7 recognizes the report, but has no PIM state for this group.
- CE-7 sends a PIM join toward the RP, in this case the nearest RP is PE-3. PE-3 already has (S,G) state for this group, so forwards traffic toward receiver H-7.
- CE-7 does a Reverse Path Forwarding (RPF) lookup of the source address in the route table, and issues a PIM join toward the source.

The join is propagated across the provider network toward PE-1, which is the resolved RPF next hop for the source.

```
*A:CE-7# show router 1 pim group type sg detail
=====
PIM Source Group ipv4
=====
Group Address       : 225.0.0.1
Source Address      : 192.168.55.2
RP Address          : 10.2.3.5
Advt Router        :
Flags               : spt                Type           : (S,G)
Mode                : sparse
MRIB Next Hop      : 172.16.37.1
MRIB Src Flags     : remote
Keepalive Timer Exp: 0d 00:02:37
Up Time            : 0d 00:00:54          Resolved By    : rtable-u

Up JP State        : Joined              Up JP Expiry   : 0d 00:00:05
Up JP Rpt         : Not Pruned          Up JP Rpt Override : 0d 00:00:00

Register State     : No Info
Reg From Anycast RP: No

Rpf Neighbor       : 172.16.37.1
```

```

Incoming Intf      : int-CE-7-PE-3
Outgoing Intf List : int-CE-7-H-7

Curr Fwding Rate   : 4811.800 kbps
Forwarded Packets  : 32917
Forwarded Octets   : 32324494
Spt threshold      : 0 kbps
Admin bandwidth    : 1 kbps
Discarded Packets  : 0
RPF Mismatches     : 0
ECMP opt threshold : 7
-----
Groups : 1
=====
    
```

The join is received by CE-5, which contains the subnet of the source.

CE-5 now recognizes the multicast group as a valid stream. CE-5 becomes the root of the shortest path tree for the group.

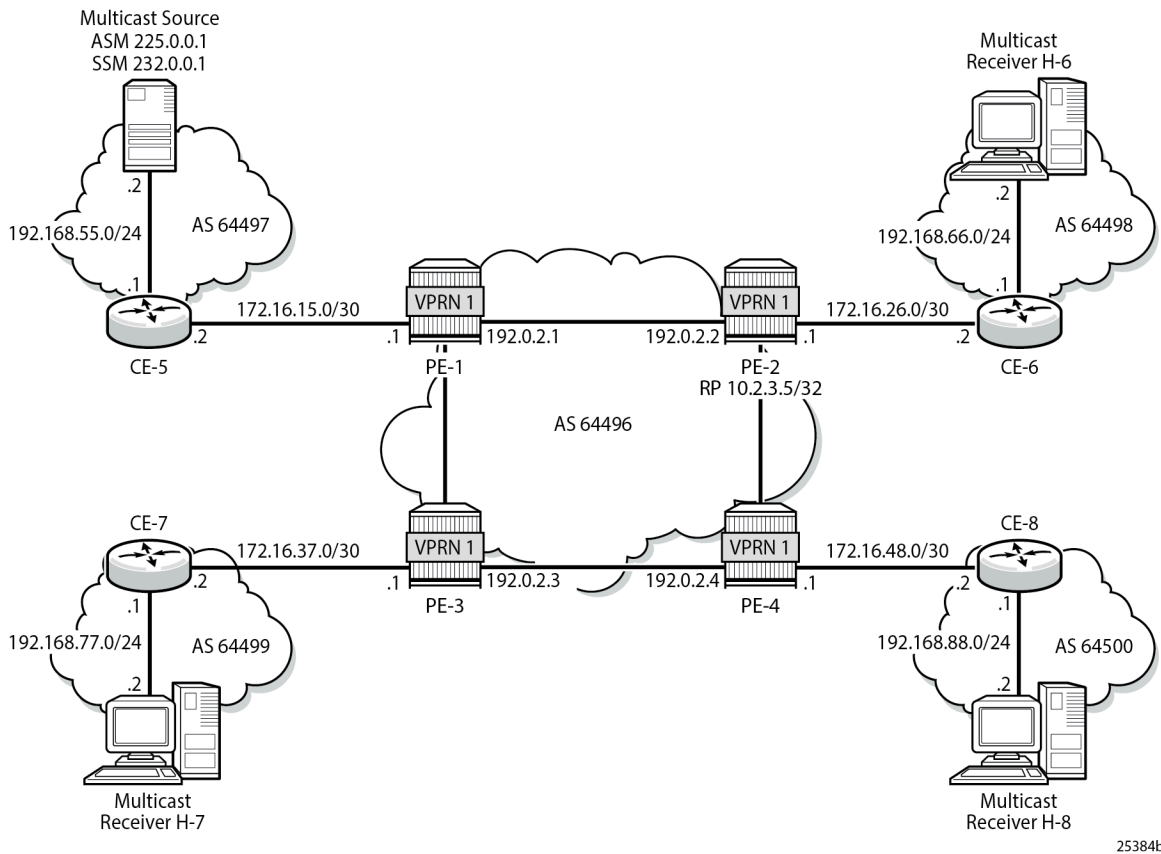
```

A:CE-5# show router 1 pim group
=====
Legend:  A = Active   S = Standby
=====
PIM Groups ipv4
=====
Group Address          Type          Spt Bit  Inc Intf      No.0ifs
  Source Address       RP
-----
225.0.0.1              (S,G)        spt      int-CE-5-S-5  1
  192.168.55.2         10.2.3.5
-----
Groups : 1
=====
    
```

## PIM Source Specific Multicasting

There is no requirement for an RP, because customer multicast signaling is PIM-SSM. The multicast group address used for the PMSI must be the same on all PEs for this VPRN instance.

Figure 37: PIM SSM in Customer Signaling Plane



The following shows the VPRN configuration for PIM and MVPN for PE-1.

```
On PE-1:
configure
  service
    vprn 1
      pim
        interface "int-PE-1-CE-5"
        exit
        no shutdown
      exit
    mvpn
      provider-tunnel
        inclusive
        pim asm 239.255.255.1
      exit
    exit
  exit
  no shutdown
exit all
```

There is a similar configuration required for each of the other PEs. Verify that PIM in the GRT has signaled the I-PMSIs.

For the PE acting as the RP for global PIM:

```
A:PE-1# show router pim group
=====
Legend:  A = Active  S = Standby
=====
PIM Groups ipv4
=====
Group Address          Type          Spt Bit  Inc Intf      No.0ifs
  Source Address        RP           State    Inc Intf(S)
-----
239.255.255.1         (*,G)                3
*                    192.0.2.1
239.255.255.1         (S,G)          spt      system        3
192.0.2.1             192.0.2.1
239.255.255.1         (S,G)          spt      int-PE-1-PE-2 3
192.0.2.2             192.0.2.1
239.255.255.1         (S,G)          spt      int-PE-1-PE-3 3
192.0.2.3             192.0.2.1
239.255.255.1         (S,G)          spt      int-PE-1-PE-2 2
192.0.2.4             192.0.2.1
-----
Groups : 5
=====
```

PE-3 has:

```
A:PE-3# show router pim group
=====
Legend:  A = Active  S = Standby
=====
PIM Groups ipv4
=====
Group Address          Type          Spt Bit  Inc Intf      No.0ifs
  Source Address        RP           State    Inc Intf(S)
-----
239.255.255.1         (*,G)                1
*                    192.0.2.1
239.255.255.1         (S,G)          spt      system        2
192.0.2.3             192.0.2.1
-----
Groups : 2
=====
```

This shows a (S,G) join toward the RP at 192.0.2.1, plus a (\*,G) join from RP. These represent the outgoing and incoming PIM interfaces for the VRF.

This results in a series of PIM neighbors through the I-PMSIs within the VRF, which are maintained using PIM hellos.

```
A:PE-1# show router 1 pim neighbor
=====
PIM Neighbor ipv4
=====
Interface          Nbr DR Prty   Up Time      Expiry Time  Hold Time
  Nbr Address
-----
int-PE-1-CE-5     1             0d 00:21:55  0d 00:01:21  105
172.16.15.2
```

```

1-mt-239.255.255.1      1          0d 00:22:30   0d 00:01:18   105
  192.0.2.2
1-mt-239.255.255.1      1          0d 00:22:17   0d 00:01:31   105
  192.0.2.3
1-mt-239.255.255.1      1          0d 00:22:06   0d 00:01:42   105
  192.0.2.4
-----
Neighbors : 4
=====

```

## PIM SSM - Customer Edge Router Multicast Configuration

Each CE router has a PIM neighbor peer relationship with its nearest PE.

The CE router (CE-5) containing the source has PIM enabled on the interface connected to the source.

```

On CE-5:
configure
  service
    vprn 1
      pim
        interface "int-CE-5-PE-1"
        exit
        interface "int-CE-5-S-5"
        exit
        no shutdown
      exit
    no shutdown
  exit all

```

The CE containing the receivers has IGMP enabled on the interface connected to the receivers and PIM on the interface facing the PE.

```

On CE-6:
configure
  service
    vprn 1
      igmp
        interface "int-CE-6-H-6"
        exit
      exit
      pim
        interface "int-CE-6-PE-2"
        exit
        no shutdown
      exit
      static-route-entry 192.168.55.0/24
        next-hop 172.16.26.1
        no shutdown
      exit
    no shutdown
  exit all

```

## Traffic Flow

The source multicasts a stream with group address 232.0.0.1 toward CE-5. When there is no receiver interested in the group at this time, there are no outgoing interfaces, so the Outgoing Interface List (OIL) is empty.

```
A:CE-5# show router 1 pim group
=====
Legend:  A = Active   S = Standby
=====
PIM Groups ipv4
=====
Group Address          Type          Spt Bit  Inc Intf      No.Oifs
Source Address         RP           State    Inc Intf(S)
-----
---snip---
232.0.0.1              (S,G)                int-CE-5-S-5   0
192.168.55.2
-----
Groups : 2
=====
```

The receiver H-6 wants to join the group 232.0.0.1, and so sends in an IGMPv3 report toward CE-6. CE-6 recognizes the report, which contains the source 192.168.55.2 in the include filter list.

```
*A:CE-6# show router 1 igmp group
=====
IGMP Interface Groups
=====
(192.168.55.2,232.0.0.1)          UpTime: 0d 00:00:58
  Fwd List : int-CE-6-H-6
-----
Entries : 1
=====
IGMP Host Groups
=====
No Matching Entries
=====
IGMP SAP Groups
=====
No Matching Entries
=====
IGMP SLA Profile Instance Groups
=====
No Matching Entries
=====
```

CE-6 does a RPF lookup of the source address in the route table, and issues a PIM join toward the source. The join is propagated across the provider network, toward PE-1 which is the resolved RPF next hop for the source.

```
A:PE-1# show router 1 pim group detail
=====
PIM Source Group ipv4
=====
Group Address      : 232.0.0.1
```

```

Source Address      : 192.168.55.2
RP Address          : 0
Advt Router        : 172.16.15.2
Flags              :                               Type           : (S,G)
Mode               : sparse
MRIB Next Hop     : 172.16.15.2
MRIB Src Flags    : remote
Keepalive Timer   : Not Running
Up Time           : 0d 00:01:08           Resolved By          : rtable-u

Up JP State        : Joined                Up JP Expiry         : 0d 00:00:51
Up JP Rpt         : Not Joined StarG      Up JP Rpt Override  : 0d 00:00:00

Register State    : No Info
Reg From Anycast RP: No

Rpf Neighbor      : 172.16.15.2
Incoming Intf     : int-PE-1-CE-5
Outgoing Intf List : 1-mt-239.255.255.1

Curr Fwding Rate  : 4811.800 kbps
Forwarded Packets : 41936                  Discarded Packets   : 0
Forwarded Octets  : 41181152              RPF Mismatches     : 0
Spt threshold    : 0 kbps                  ECMP opt threshold : 7
Admin bandwidth  : 1 kbps

-----
Groups : 1
=====

```

The outgoing interface is the I-PMSI: 1-mt-239.255.255.1.

The join is received by CE-5, which contains the subnet of the source.

CE-5 now recognizes the multicast group as a valid stream. CE-5 becomes the root of the shortest path tree for the group.

```

A:CE-5# show router 1 pim group

=====
Legend:  A = Active   S = Standby
=====
PIM Groups ipv4
=====
Group Address      Type           Spt Bit  Inc Intf      No.Oifs
Source Address     RP            State    Inc Intf(S)
-----
232.0.0.1          (S,G)                int-CE-5-S-5  1
192.168.55.2

-----
Groups : 1
=====

```

## PE BGP Auto-Discovery

Discovery of Multicast-enabled Virtual Private Networks (MVPNs) can also be achieved using BGP. To this end, any PE that is a member of a multicast VPN advertises this using a BGP multi-protocol Network Layer Reachability Information (NLRI) update that is sent to all PEs within the AS. This update contains an intra-AS I-PMSI Auto-Discovery route type, also known as an Intra-AD. These use a dedicated address family

**mvpn-ipv4** so each PE must be configured to originate and accept such updates. The following needs to be modified in the **bgp** context for all PE nodes:

```
On all PEs:
configure
router
  bgp
    group "INTERNAL"
      family vpn-ipv4 mvpn-ipv4
    exit all
```

This is achieved in the GRT within the **bgp** context.

This allows each BGP speaker to advertise its capabilities within a BGP Open message.

The following BGP summary on PE-1 shows that BGP sessions are established between the PEs for address families VPN-IPv4 and MVPN-IPv4 in the base routing instance:

```
*A:PE-1# 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
-----
192.0.2.2
Def. Inst          64496      4   0 00h00m26s 0/0/0 (VpnIPv4)
                   6   0      0/0/0 (MvpnIPv4)
192.0.2.3
Def. Inst          64496      4   0 00h00m24s 0/0/0 (VpnIPv4)
                   6   0      0/0/0 (MvpnIPv4)
192.0.2.4
Def. Inst          64496      4   0 00h00m23s 0/0/0 (VpnIPv4)
                   6   0      0/0/0 (MvpnIPv4)
172.16.15.2
1                  64497     86   0 00h38m27s 1/1/1 (IPv4)
                   86   0
-----
```

## BGP Auto-Discovery - PE VPRN Multicast Configuration

Each PE contains a CE which is part of the multicast VRF, so it is necessary to enable PIM on each interface containing an attachment circuit toward a CE, and to configure the I-PMSI multicast tunnel for the VRF.

For the BGP routes to be accepted into the VRF, a route-target community is required (vrf-target). This is configured in the **configure service vprn 1 mvpn** context and, in this case, is set to the same value as the unicast vrf-target, the vrf-target community in the **configure service vprn 1 bgp-ipvpn mpls vrf-target** context.

On each PE, the **mvpn** context of the VPRN instance is configured as follows:

```
On PE-2:
```



```

configure
  service
    vprn 1
      mvpn
        auto-discovery
        provider-tunnel
          inclusive
            pim asm 239.255.255.1
          exit
        exit
      exit
    vrf-target unicast
    exit
  no shutdown
exit all

```

The multicast group address used for the PMSI must be the same on all PEs for this VPRN instance.

The presence of auto-discovery initiates BGP updates between the PEs that contain an MVPN, such as Intra-AD MVPN routes, are generated and advertised to each peer

```

*A:PE-1# show router bgp routes mvpn-ipv4
=====
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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP MVPN-IPv4 Routes
=====
Flag  RouteType      OriginatorIP      LocalPref  MED
      RD             SourceAS          Path-Id     IGP Cost
      Nexthop        SourceIP          Label
      As-Path        GroupIP
-----
u*>i  Intra-Ad          192.0.2.2        100        0
      64496:1        -                None       -
      192.0.2.2     -
      No As-Path    -
u*>i  Intra-Ad          192.0.2.3        100        0
      64496:1        -                None       -
      192.0.2.3     -
      No As-Path    -
u*>i  Intra-Ad          192.0.2.4        100        0
      64496:1        -                None       -
      192.0.2.4     -
      No As-Path    -
-----
Routes : 3
=====

```

This shows that PE-1 has received an Intra-AD route from each of the other PEs, each of which has multicast VPRN 1 configured.

Examining the intra-AD routes received from PE-2 shows that the route-target community matches the unicast VRF-target (64496:1), and also that the PMSI tree has a multicast group address of 239.255.255.1, which matches the I-PMSI group configuration on PE-1.

```
*A:PE-1# show router bgp routes mvpn-ipv4 type intra-ad originator-ip 192.0.2.2 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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP MVPN-IPv4 Routes
=====
Original Attributes

Route Type      : Intra-Ad
Route Dist.     : 64496:1
Originator IP   : 192.0.2.2
Nexthop        : 192.0.2.2
Path Id         : None
From           : 192.0.2.2
Res. Nexthop    : 0.0.0.0
Local Pref.     : 100
Aggregator AS  : None
Atomic Aggr.   : Not Atomic
AIGP Metric     : None
Connector      : None
Community       : no-export target:64496:1
Cluster        : No Cluster Members
Originator Id   : None
Peer Router Id  : 192.0.2.2
Flags          : Used Valid Best IGP
Route Source    : Internal
AS-Path        : No As-Path
Route Tag      : 0
Neighbor-AS    : n/a
DB Orig Val    : N/A
Source Class    : 0
Add Paths Send : Default
Last Modified  : 00h01m39s
VPRN Imported  : 1
Interface Name  : NotAvailable
Aggregator     : None
MED            : 0
IGP Cost       : n/a
Final Orig Val : N/A
Dest Class     : 0
-----
PMSI Tunnel Attributes :
Tunnel-type      : PIM-SM Tree
Flags             : Type: RNVE(0) BM: 0 U: 0 Leaf: not required
MPLS Label       : 0
Sender           : 192.0.2.2
P-Group         : 239.255.255.1
-----
---snip---
-----
Routes : 1
=====
```

Verify that PIM in the GRT has signaled the I-PMSIs.

For the PE acting as the RP for global PIM:

```
*A:PE-1# show router pim group
=====
```

```

Legend:  A = Active   S = Standby
=====
PIM Groups ipv4
=====
Group Address          Type          Spt Bit   Inc Intf   No.0ifs
  Source Address       RP           State     Inc Intf(S)
-----
239.255.255.1         (*,G)                3
*                      192.0.2.1
239.255.255.1         (S,G)          spt       system     3
  192.0.2.1           192.0.2.1
239.255.255.1         (S,G)          spt       int-PE-1-PE-2  3
  192.0.2.2           192.0.2.1
239.255.255.1         (S,G)          spt       int-PE-1-PE-3  3
  192.0.2.3           192.0.2.1
239.255.255.1         (S,G)          spt       int-PE-1-PE-2  2
  192.0.2.4           192.0.2.1
-----
Groups : 5
=====

```

This shows an incoming (S,G) join from all other PEs within the multicast VRF, plus an outgoing (\*,G) join to the same PEs.

PE-3 has the following PIM groups:

```

*A:PE-3# show router pim group
=====
Legend:  A = Active   S = Standby
=====
PIM Groups ipv4
=====
Group Address          Type          Spt Bit   Inc Intf   No.0ifs
  Source Address       RP           State     Inc Intf(S)
-----
239.255.255.1         (*,G)                1
*                      192.0.2.1
239.255.255.1         (S,G)          spt       system     2
  192.0.2.3           192.0.2.1
-----
Groups : 2
=====

```

This shows a (S,G) join toward the RP at 192.0.2.1, plus a (\*,G) join from RP. These represent the outgoing and incoming PIM interfaces for the VRF.

This results in a series of PIM neighbors through the I-PMSIs within the VRF. The neighbors were discovered using BGP (not with PIM as per Rosen MVPN), therefore, there are no PIM hellos exchanged.

```

*A:PE-1# show router 1 pim neighbor
=====
PIM Neighbor ipv4
=====
Interface          Nbr DR Prty   Up Time      Expiry Time  Hold Time
  Nbr Address
-----
int-PE-1-CE-5      1              0d 00:32:28  0d 00:01:18  105
  172.16.15.2
1-mt-239.255.255.1 1              0d 00:01:53  never        65535
  192.0.2.2

```

```

1-mt-239.255.255.1      1          0d 00:01:53  never      65535
 192.0.2.3
1-mt-239.255.255.1      1          0d 00:01:52  never      65535
 192.0.2.4
-----
Neighbors : 4
=====

```

## BGP Auto-Discovery - Customer Signaling Domain

The customer signaling is independent from the provider PE discovery mechanism, therefore, all of the customer signaling techniques described when using PIM for auto-discovery within provider domain are also applicable when using BGP for auto-discovery, namely

- PIM Any Source Multicasting with RP at the provider PE
- PIM Any Source Multicasting with Anycast RP at the provider PE
- PIM Source Specific Multicasting

## Data Path Using Selective PMSI

When a configurable data threshold for a multicast group has been exceeded, multicast traffic across the provider network can be switched to a Selective PMSI (S-PMSI).

This has to be configured as a separate group and must contain a threshold which, if exceeded, sees a new PMSI signaled by the PE nearest the source, and traffic switched onto the S-PMSI.

```

On PE-1:
configure
  service
    vprn 1
      mvpn
        provider-tunnel
          inclusive
            pim asm 239.255.255.1
          exit
        exit
        selective
          data-threshold 232.0.0.0/8 1
          pim-ssm 232.255.1.0/24
        exit
      exit
    exit
  no shutdown
exit all

```

This shows that when the traffic threshold for multicast groups covered by the range 232.0.0.0/8 exceeds 1 kb/s between a pair of PEs, then an S-PMSI is signaled between the PEs. This is a separate multicast tunnel over which traffic in that group now flows.

```

*A:PE-1# show router 1 pim s-pmsi detail

=====
PIM Selective provider tunnels
=====
Md Source Address   : 192.0.2.1           Md Group Address   : 232.255.1.0
Number of VPN SGs  : 1                Uptime             : 0d 00:00:59

```

```

MT IfIndex      : 5798468
Multistream-Id  : 0

VPN Group Address : 232.0.0.1
VPN Source Address : 192.168.55.2
State : TX Joined           Mdt Threshold : 1
Join Timer      : 0d 00:00:01   Holddown Timer : 0d 00:00:00
=====
PIM Selective provider tunnels Interfaces : 1
=====
    
```

In this example, the (S,G) group is (192.168.55.2, 232.0.0.1). When the data rate has exceeded the configured MDT threshold of 1 kb/s, a new provider tunnel with a group address of 232.255.1.0 has been signaled and now carries the multicast stream.

The TX Joined state indicates that the S-PMSI has been sourced at this PE -PE-1.

Comparing this to PE-3, where a receiver is connected through a CE indicates that it has received a join to connect the S-PMSI.

```

*A:PE-3# show router 1 pim s-pmsi detail

=====
PIM Selective provider tunnels
=====
Md Source Address : 192.0.2.1           Md Group Address : 232.255.1.0
Number of VPN SGs : 1                   Uptime           : 0d 00:01:02
MT IfIndex        : 5806657             Egress Fwding Rate : 4815.728 kbps
Multistream-Id    : 0

VPN Group Address : 232.0.0.1
VPN Source Address : 192.168.55.2
State : RX Joined
Expiry Timer      : 0d 00:02:58
=====
PIM Selective provider tunnels Interfaces : 1
=====
    
```

## Conclusion

This chapter provides configuration on how to configure multicast within a VPRN with NG MVPN techniques. Specifically, discovery of multicast VPNs using PIM and BGP auto-discovery mechanisms are described with a number of ASM and SSM signaling techniques within the customer domain.

# NG-MVPN Inter-AS Model B Using Non-Segmented mLDP Tunnels

This chapter provides information about NG-MVPN Inter-AS Model B Using Non-Segmented mLDP Tunnels.

Topics in this chapter include:

- [Applicability](#)
- [Overview](#)
- [Configuration](#)
- [Conclusion](#)

## Applicability

The information and configuration in this chapter are based on SR OS Release 23.10.R2.

There is no specific configuration required to support non-segmented mLDP for inter-AS model B. However, VPN-recursive mLDP Forwarding Equivalence Class (FEC) functionality must be supported.

The configuration of multicast in a VPRN is described in the chapter [NG-MVPN Configuration with MPLS](#).

## Overview

Multicast in an inter-AS model B network using Draft-Rosen techniques is described in the chapter [Rosen MVPN Inter-AS Option B](#), where the set of Multicast Distribution Trees (MDTs) are signaled using Protocol Independent Multicast Source-Specific Mode (PIM-SSM).

It is also possible to create MDTs between PEs in different Autonomous Systems (ASs) for Next-Generation MVPN (NG-MVPN) using non-segmented dynamic multicast LDP (mLDP) trees, where the root and leaf PEs are in different ASs. This chapter describes the configuration of MVPN services between PEs in different ASs using NG-MVPN techniques.

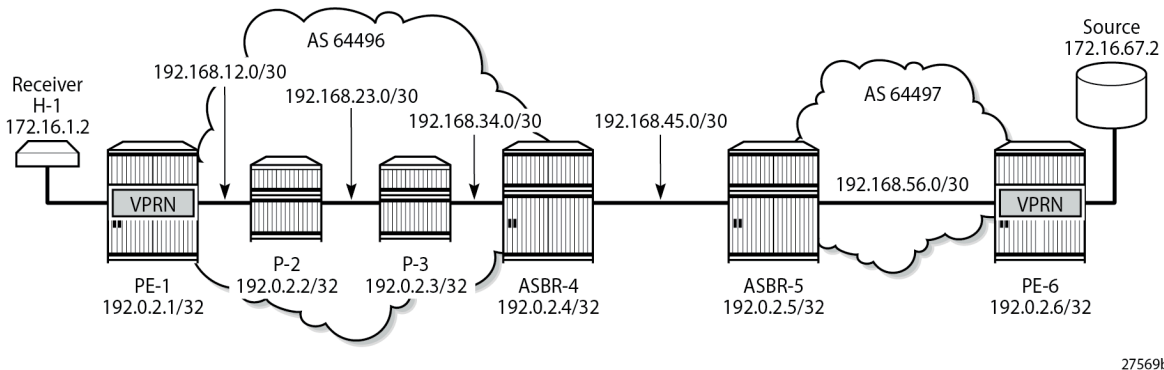
**Figure 38: NG-MVPN Inter-AS Model B** shows an example of a network comprising routers in two neighboring ASs, modeled as an inter-AS model B network.

A VPRN instance exists on PE-1 and PE-6, with a single source and receiver. PE-1 is connected to its local Autonomous System Border Router (ASBR) via a pair of core routers, P-2 and P-3, neither of which are members of the VPRN instance. P-2 acts as a Route Reflector (RR) for AS 64496.

The source router generates multicast traffic with a group address of 239.255.0.1, and the receiver H-1 in AS 64496 becomes a member of the multicast group using IGMPv3 signaling.

No multicast protocols are configured on the ASBRs.

Figure 38: NG-MVPN Inter-AS Model B



In an MVPN, a transport tunnel is signaled that carries multicast traffic from source PE to any receiver PE that contains attached multicast routers, or hosts that want to become multicast group members. In this case, a multicast LDP (mLDP) tunnel is signaled between source PE and destination PEs to carry multicast traffic across the multi-AS provider network.

The mLDP Provider Multicast Service Interface (PMSI), also known as the provider tunnel, is established between PEs that declare membership of the MVPN, by generating and advertising an MVPN type 1 intra-AD BGP route. This route contains a PMSI Tunnel Attribute (PTA) that describes the tunnel type, the root node, and the LSP ID. Upon receipt of the intra-AD route, the receiving PE checks that the route is valid and can be imported into the VPRN instance. If the route is valid, the receiving router signals a point-to-multipoint (P2MP) LDP label mapping message toward the root address contained within the PTA of the intra-AD route. At the root, MPLS-encapsulated multicast traffic is forwarded to the downstream router by pushing on a label received from the downstream router.

Inter-AS model B unicast has VPN-IPv4 routes advertised from one AS to the other across the AS boundary. No system addresses of the PEs within an AS are advertised across the AS boundary, so the path for inter-AS unicast traffic is resolved using the labeled VPN-IPv4 routes via the ASBR, using MPLS encapsulation. In a unicast environment, traffic from PE-1 to PE-6 would be encapsulated in a tunnel to ASBR-4, de-encapsulated at ASBR-4, forwarded to ASBR-5, and on toward PE-6. The tunnel comprises the MPLS transport label plus the label associated with the VPN-IPv4 route.

For unicast routes, BGP next-hop-self is performed on the ASBR (from a control plane perspective) while service labels are swapped at the ASBR within the data plane side. This results in a segmented approach.

Multicast traffic requires a non-segmented provider tunnel to be routed from the root PE toward the receiver routers. This means that the tunnel must traverse the AS boundary without de-encapsulation, and therefore, must be non-segmented.

If the provider tunnel uses mLDP, the receivers initiate the signaling by sending an LDP label mapping message along the control path toward the root. This follows the path of the intra-AD route that advertises the root of the I-PMSI. An mLDP label mapping message consists of an allocated label L, with FEC element <X,Y>, where X identifies the root node and Y is the opaque value, so the P2MP label mapping can be denoted as <X,Y,L>.

The FEC element contains the root address of the LSP plus a variable length opaque value. The opaque value contains information meaningful to the root and leaf routers, but not to intermediate routers; for example, a P2MP LSP-ID or a nested opaque value.

The root address is the system address of the router that advertised the intra-AD route. PE-1 has no unicast route to PE-6, but has learned the ASBR-4 address from the BGP next-hop of the intra-AD MVPN route.

However, in an mLDP environment, each router must take part in the signaling of the P2MP LSP, but not every router has an MVPN route, so any mLDP label mapping message received by P-3 to the root address of PE-6 is dropped.

A solution to this is described in RFC 6512, *Using Multipoint LDP When the Backbone Has No Route to the Root*.

PE-1 signals an mLDP LSP as if the root is at ASBR-4. P-2 and P-3 have a route to ASBR-4 as they are part of the same IGP instance. The actual root address on PE-6 is encapsulated within the mLDP label mapping message originated by PE-1 as an inner root address. This is a recursive FEC type, where the actual root FEC element is encapsulated within a FEC element as an opaque value that has a root at the ASBR.

ASBR-4 does not have a unicast route toward PE-6, but it has received the intra-AD MVPN route advertised by PE-6. This intra-AD route contains the BGP next-hop of ASBR-5, so a path toward PE-6 exists if the address in this route is used. To distinguish between any number of intra-AD routes at the ASBR, the recursive FEC contains the intra-AD route-distinguisher (RD) as an opaque value, which is used with the root address to match the correct intra-AD route.

This recursive FEC is defined as a VPN-recursive FEC, because the VPN intra-AD route is used as the route lookup to forward the label mapping message.

[Table 2: mLDP Message Opaque Value Types in MVPN Model B](#) shows the opaque value types used in MVPN model B.

*Table 2: mLDP Message Opaque Value Types in MVPN Model B*

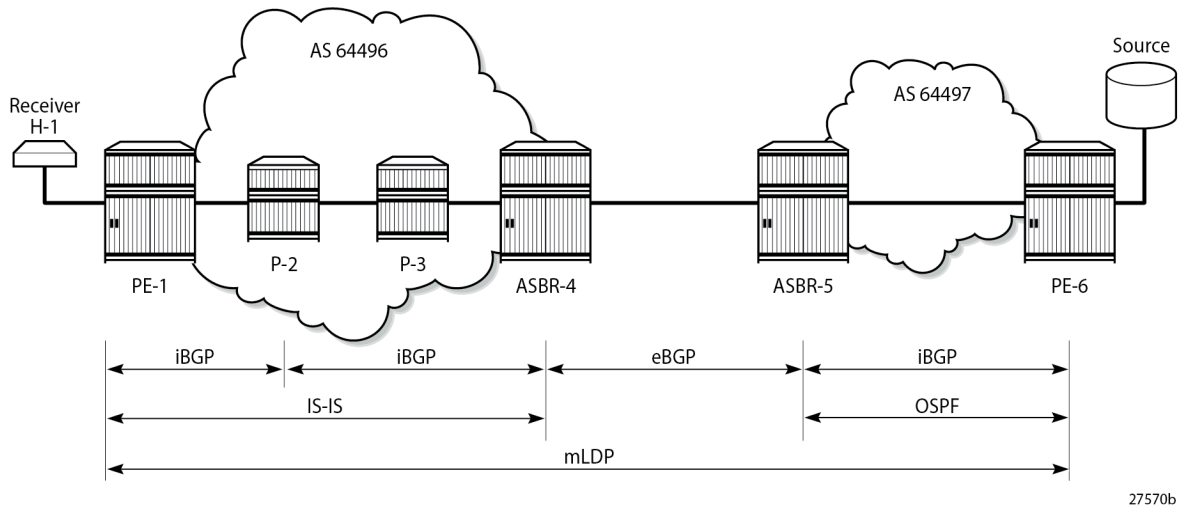
Opaque Type	Opaque Name	Use	FEC Element Representation
1	Generic	VPRN local AS	Root, Opaque<P2MP ID>
8	VPN Recursive	Inter-AS model B mLDP	<ASBR, Opaque <RD, Root, P2MP ID>>

## Configuration

[Figure 39: Inter-AS MVPN Protocol Requirements](#) shows the required protocol configuration and peering.



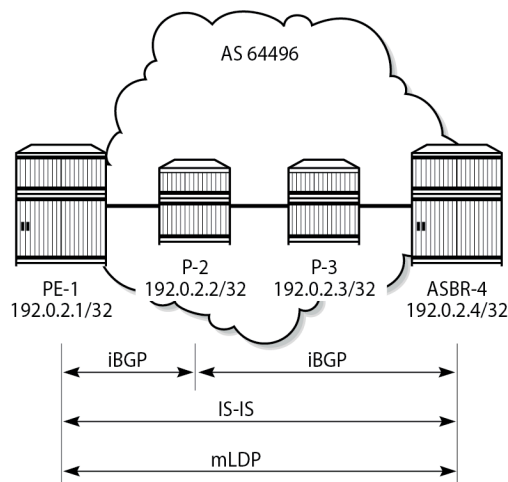
Figure 39: Inter-AS MVPN Protocol Requirements



## AS 64496

Figure 40: AS 64496 Protocols shows the protocol requirements for AS 64496.

Figure 40: AS 64496 Protocols



## Router Interface and IS-IS Configuration

The first step is to configure the router interfaces and IS-IS as the Interior Gateway Protocol (IGP) in AS 64496.

The router interfaces for PE-1 are configured as follows:

```
# on PE-1:
configure
router
  interface "int-PE-1-P-2"
    address 192.168.12.1/30
    port 1/1/c1/1
  exit
  interface "system"
    address 192.0.2.1/32
  exit
```

Each interface is configured to run IS-IS as the IGP. Each router is configured as a level 2 router.

```
# on PE-1:
configure
router
  isis 0
    level-capability level-2
    area-id 49.0001
    traffic-engineering
    level 2
      wide-metrics-only
    exit
    interface "system"
      level-capability level-2
    exit
    interface "int-PE-1-P-2"
      level-capability level-2
      interface-type point-to-point
    exit
    no shutdown
  exit
```

The configuration for all other nodes in the AS is the same, apart from the IP addresses. The IP addresses can be derived from [Figure 38: NG-MVPN Inter-AS Model B](#).

## LDP Configuration

Label Distribution Protocol (LDP) is used as the MPLS protocol and must be enabled on each router interface, as follows:

```
# on PE-1:
configure
router
  ldp
    import-pmsi-routes
      mvpn # needed for bindings
    exit
    tcp-session-parameters
    exit
    interface-parameters
      interface "int-PE-1-P-2" dual-stack
        ipv4
          fec-type-capability
          p2mp-ipv4 enable # default
        exit
      exit
    exit
```

```

        exit
    exit
targeted-session
    exit
exit

```

This configuration must be repeated on each of the other routers in the AS. As LDP is used as the provider tunnel interface for multicast traffic, each interface must also support P2MP LDP tunnels. Therefore, the FEC type capability for IPv4 P2MP tunnels must be enabled. The default value is enable, but is included in the preceding configuration for clarity.

## BGP Configuration

### P-2 Route Reflector

P-2 is configured as an RR within AS 64496 and peers with both PE-1 and PE-4. The address families negotiated are VPN-IPv4 for unicast VPRN routes, and MVPN-IPv4 routes for multicast VPRN routes. The cluster ID is set to ensure that P-2 is an RR.

```

# on P-2:
configure
router
    autonomous-system 64496
    bgp
        cluster 192.0.2.2
        group "internal"
            family vpn-ipv4 mvpn-ipv4
            type internal
            cluster 192.0.2.2
            neighbor 192.0.2.1
            exit
            neighbor 192.0.2.4
            exit
        exit
    no shutdown
exit

```

### PE-1

PE-1 is a BGP peer of RR P-2, as follows:

```

# on PE-1:
configure
router
    autonomous-system 64496
    bgp
        group "internal"
            family vpn-ipv4 mvpn-ipv4
            type internal
            neighbor 192.0.2.2
            exit
        exit
    no shutdown
exit

```

## ASBR-4

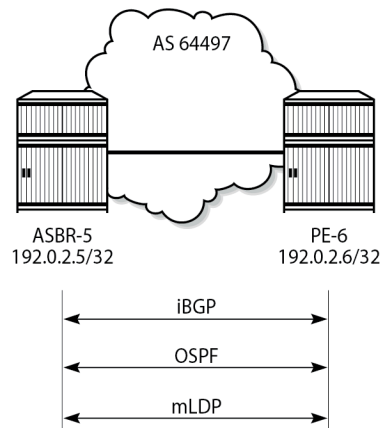
For completeness, the ASBR-4 BGP configuration is as follows.

```
# on ASBR-4:
configure
router
  autonomous-system 64496
  bgp
    family vpn-ipv4 mvpn-ipv4
    group "internal"
      type internal
      neighbor 192.0.2.2
    exit
  exit
  no shutdown
exit
```

## AS 64497

[Figure 41: AS 64497 Protocols](#) shows the protocol requirements for AS 64497.

*Figure 41: AS 64497 Protocols*



27572b

## Router Interface and OSPF Configuration

The first step is to configure router interfaces and OSPF on each router shown in [Figure 41: AS 64497 Protocols](#). All router interfaces are members of a single backbone area: area 0.0.0.0.

The following router interfaces are configured on PE-6:

```
# on PE-6:
configure
router
  interface "int-PE-6-ASBR-5"
    address 192.168.56.2/30
    port 1/1/c2/1
```

```

exit
interface "system"
  address 192.0.2.6/32
exit

```

The configuration for PE-6 to enable OSPF is:

```

# on PE-6:
configure
  router
    ospf 0
      area 0.0.0.0
        interface "int-PE-6-ASBR-5"
          interface-type point-to-point
        exit
      interface "system"
        interface-type point-to-point
      exit
    exit
  no shutdown
exit

```

## LDP Configuration

LDP is used as the MPLS protocol and must be enabled on each router interface, as follows.

```

# on PE-6:
configure
  router
    ldp
      import-pmsi-routes
      exit
      tcp-session-parameters
      exit
      interface-parameters
        interface "int-PE-6-ASBR-5" dual-stack
          ipv4
            fec-type-capability
            p2mp-ipv4 enable
          exit
        exit
      exit
    exit
  targeted-session
  exit
exit

```

This configuration must be repeated on each of the other routers in the AS. Again, the default value of FEC type capability for P2MP is enable, but is included for clarity.

## BGP Configuration

Within AS 64497, internal BGP peering is required between ASBR-5 and PE-6 for the VPN-IPv4 and MVPN-IPv4 address families.

The following shows the BGP configuration for such a peering.

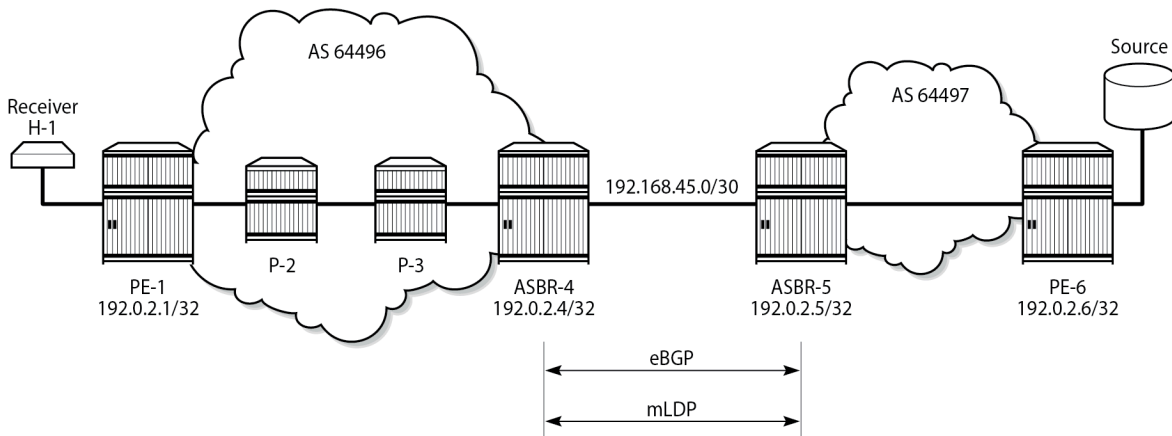
```
# on ASBR-5:
configure
router
  autonomous-system 64497
  bgp
    family vpn-ipv4 mvpn-ipv4
    group "internal"
      type internal
      neighbor 192.0.2.6
    exit
  exit
  no shutdown
exit
```

```
# on PE-6:
configure
router
  autonomous-system 64497
  bgp
    rapid-withdrawal
    rapid-update mvpn-ipv4 vpn-ipv4
    group "internal"
      family vpn-ipv4 mvpn-ipv4
      type internal
      neighbor 192.0.2.5
    exit
  exit
  no shutdown
exit
```

## Inter-AS Configuration

Figure 42: [Inter-AS Protocols](#) shows the protocols required between ASBR-4 and ASBR-5. The LDP transport address and the BGP speaker peer addresses are the interface addresses.

Figure 42: *Inter-AS Protocols*



27573b

## eBGP Peering

The following shows the eBGP peering configuration for ASBR-4. The peer address is the interface address of ASBR-5.

```
# on ASBR-4:
configure
router
  autonomous-system 64496
  bgp
    family vpn-ipv4 mvpn-ipv4
    enable-inter-as-vpn          # required for inter-AS VPN option B
    split-horizon
    group "external"
      peer-as 64497
      neighbor 192.168.45.2
    exit
  exit
  no shutdown
exit
```

Similarly, the BGP configuration for ASBR-5 peering toward ASBR-4 is as follows:

```
# on ASBR-5:
configure
router
  autonomous-system 64497
  bgp
    family vpn-ipv4 mvpn-ipv4
    enable-inter-as-vpn          # required for inter-AS VPN option B
    split-horizon
    group "external"
      peer-as 64496
      neighbor 192.168.45.1
    exit
  exit
  no shutdown
exit
```

Verification of the BGP peering session between ASBR-4 and ASBR-5 is shown in the following output:

```
# on ASBR-4:
*A:ASBR-4# show router bgp summary group "external"
=====
BGP Router ID:192.0.2.4      AS:64496      Local AS:64496
=====
BGP Admin State      : Up      BGP Oper State      : Up
Total Peers          : 1
Total IPv4 Remote Rts : 0      Total IPv4 Rem. Active Rts : 0
Total IPv6 Remote Rts : 0      Total IPv6 Rem. Active Rts : 0
Total IPv4 Backup Rts : 0      Total IPv6 Backup Rts    : 0
Total LblIPv4 Rem Rts : 0      Total LblIPv4 Rem. Act Rts : 0
Total LblIPv6 Rem Rts : 0      Total LblIPv6 Rem. Act Rts : 0
Total LblIPv4 Bkp Rts : 0      Total LblIPv6 Bkp Rts    : 0

Total VPN-IPv4 Rem. Rts : 0      Total VPN-IPv4 Rem. Act. Rts: 0
Total VPN-IPv6 Rem. Rts : 0      Total VPN-IPv6 Rem. Act. Rts: 0
Total VPN-IPv4 Bkup Rts : 0      Total VPN-IPv6 Bkup Rts   : 0
```

```

Total MVPN-IPv4 Rem Rts : 0      Total MVPN-IPv4 Rem Act Rts : 0
Total MVPN-IPv6 Rem Rts : 0      Total MVPN-IPv6 Rem Act Rts : 0
Total MDT-SAFI Rem Rts : 0       Total MDT-SAFI Rem Act Rts : 0
Total McIPv4 Remote Rts : 0      Total McIPv4 Rem. Active Rts: 0
Total McIPv6 Remote Rts : 0      Total McIPv6 Rem. Active Rts: 0
Total McVpnIPv4 Rem Rts : 0      Total McVpnIPv4 Rem Act Rts : 0
Total McVpnIPv6 Rem Rts : 0      Total McVpnIPv6 Rem Act Rts : 0

Total EVPN Rem Rts      : 0      Total EVPN Rem Act Rts      : 0
Total L2-VPN Rem. Rts   : 0      Total L2VPN Rem. Act. Rts   : 0
Total MSPW Rem Rts     : 0      Total MSPW Rem Act Rts     : 0
Total RouteTgt Rem Rts : 0      Total RouteTgt Rem Act Rts : 0
Total FlowIpv4 Rem Rts : 0      Total FlowIpv4 Rem Act Rts : 0
Total FlowIpv6 Rem Rts : 0      Total FlowIpv6 Rem Act Rts : 0
Total FlowVpnv4 Rem Rts : 0     Total FlowVpnv4 Rem Act Rts : 0
Total FlowVpnv6 Rem Rts : 0     Total FlowVpnv6 Rem Act Rts : 0
Total Link State Rem Rts: 0      Total Link State Rem Act Rts: 0
Total SrPlcyIpv4 Rem Rts: 0     Total SrPlcyIpv4 Rem Act Rts: 0
Total SrPlcyIpv6 Rem Rts: 0     Total SrPlcyIpv6 Rem Act Rts: 0
    
```

=====

BGP Summary

=====

Legend : D - Dynamic Neighbor

=====

Neighbor

Description	AS	PktRcvd	InQ	Up/Down	State Rcv/Act/Sent (Addr Family)
		PktSent	OutQ		
-----					
192.168.45.2	64497	10	0	00h03m01s	0/0/0 (VpnIPv4)
		10	0		0/0/0 (MvpnIPv4)
-----					

## LDP Peering

LDP is configured as the MPLS protocol between ASBR-4 and ASBR-5. On ASBR-4, the interface toward ASBR-5 has LDP enabled, as follows:

```

# on ASBR-4:
configure
router
  ldp
    import-pmsi-routes
      mvpn # needed for bindings
    exit
    tcp-session-parameters
    exit
    interface-parameters
      interface "int-ASBR-4-ASBR-5" dual-stack
        ipv4
          fec-type-capability
          p2mp-ipv4 enable
        exit
        transport-address interface
      exit
    exit
  exit
  exit
  targeted-session
  exit
    
```



```
exit
```

The P2MP FEC type capability for P2MP LDP is shown. This is the default value.

For completeness, the LDP configuration on ASBR-5 for the interface toward ASBR-4 is as follows:

```
# on ASBR-5:
configure
  router
    ldp
      import-pmsi-routes
      exit
      tcp-session-parameters
      exit
      interface-parameters
        interface "int-ASBR-5-ASBR-4" dual-stack
          ipv4
            fec-type-capability
            p2mp-ipv4 enable
            exit
            transport-address interface
          exit
        exit
      exit
    exit
  targeted-session
  exit
exit
```

Verification that the LDP session is successfully established at ASBR-4 is shown in the following output:

```
# on ASBR-4:
*A:ASBR-4# show router ldp session 192.0.2.5

=====
LDP IPv4 Sessions
=====
Peer LDP Id      Adj Type  State      Msg Sent  Msg Recv  Up Time
-----
192.0.2.5:0     Link     Established  47        46        0d 00:01:35
-----
No. of IPv4 Sessions: 1
=====
---snip---
```

For completeness, the LDP session from ASBR-5 toward ASBR-4 is shown in the following output:

```
# on ASBR-5:
*A:ASBR-5# show router ldp session 192.0.2.4

=====
LDP IPv4 Sessions
=====
Peer LDP Id      Adj Type  State      Msg Sent  Msg Recv  Up Time
-----
192.0.2.4:0     Link     Established  45        47        0d 00:01:34
-----
No. of IPv4 Sessions: 1
=====
---snip---
```

When a label mapping message is received for an LDP FEC prefix, the next-hop for a FEC prefix is resolved in the routing table. The FEC is installed in the Label Information Base (LIB) if the next-hop matches a /32 route table entry.

The local interface configuration results in a route being installed with a subnet mask matching the interface configuration. In this case, the ASBR-to-ASBR route is 192.168.45.0/30.

For LDP to resolve the LDP FEC egress next-hop on ASBR-4, a /32 route matching the egress next-hop address is required in the FIB.

On ASBR-4, a static route is configured for the /32 address on ASBR-5, as follows.

```
# on ASBR-4:
configure
router
  static-route-entry 192.168.45.2/32
    next-hop 192.168.45.2
    no shutdown
  exit
exit
```

Similarly, a static route on ASBR-5 is configured for the /32 address on ASBR-4, as follows.

```
# on ASBR-5:
configure
router
  static-route-entry 192.168.45.1/32
    next-hop 192.168.45.1
    no shutdown
  exit
exit
```

The following output shows that the static route is installed in the ASBR-4 RIB.

```
# on ASBR-4:
*A:ASBR-4# show router route-table protocol static

=====
Route Table (Router: Base)
=====
Dest Prefix[Flags]                Type   Proto   Age      Pref
Next Hop[Interface Name]          Metric
-----
192.168.45.2/32                   Remote Static  00h00m00s 5
192.168.45.2                       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
=====
```

## VPRN Configuration

The VPRN service configuration for PE-1 and PE-6 is as follows:

## PE-1

```
# on PE-1:
configure
  service
    vprn 1 name "VPRN 1" customer 1 create
    interface "int-PE-1-VPRN-1-H-1" create
      address 172.16.1.1/30
      sap 1/1/c3/1:1 create
      exit
    exit
    bgp-ipvpn
      mpls
        auto-bind-tunnel
        resolution-filter
          ldp
            exit
          resolution filter
        exit
        route-distinguisher 192.0.2.1:1
        vrf-target target:64496:1
        no shutdown
      exit
    exit
    igmp
      interface "int-PE-1-VPRN-1-H-1"
        no shutdown
      exit
    no shutdown
  exit
  pim
    rp
      static
      exit
      bsr-candidate
        shutdown
      exit
      rp-candidate
        shutdown
      exit
    exit
    no shutdown
  exit
  mvpn
    auto-discovery default
    c-mcast-signaling bgp
    mdt-type receiver-only
    provider-tunnel
      inclusive
      mldp
        no shutdown
      exit
    exit
    vrf-target unicast
  exit
  no shutdown
exit
```

## PE-6

```
# on PE-6:
configure
  service
    vprn 1 name "VPRN 1" customer 1 create
    interface "int-PE-6-VPRN-1-source" create
      address 172.16.67.1/30
      sap 1/1/cl/1 create
      exit
    exit
    bgp-ipvpn
      mpls
        auto-bind-tunnel
        resolution-filter
          ldp
            exit
          resolution filter
        exit
        route-distinguisher 192.0.2.6:1
        vrf-target target:64496:1
        no shutdown
      exit
    exit
    pim
      interface "int-PE-6-VPRN-1-source"
        exit
        rp
          static
            exit
          bsr-candidate
            shutdown
          exit
          rp-candidate
            shutdown
          exit
        exit
      no shutdown
    exit
  mvpn
    auto-discovery default
    c-mcast-signaling bgp
    mdt-type sender-only
    provider-tunnel
      inclusive
      mldp
        no shutdown
      exit
    exit
    vrf-target unicast
  exit
no shutdown
exit
```

## Route Policy for MVPN Routes

The use of non-segmented LDP provider tunnels requires that Intra-AD Auto Discovery routes must be advertised across the AS boundary between PEs. Each intra-AD route generated by a PE that is a member of an MVPN contains the well-known community "No-Export", which prevents a BGP speaker from advertising the route across an AS boundary to another external BGP speaker.

In inter-AS mode B, the ASBR router must support the MVPN address family. If it receives an intra-AD route containing the No-Export community, it is not advertised to any external peer. A route policy is required to remove the No-Export community before it can be advertised across the AS boundary to a BGP speaker that has negotiated the MVPN address family capability.

In the following example configuration, the policy removes the No-Export community on PE-6, the source router from all advertised routes, by configuring the community remove action as a default action.

```
# on PE-6:
configure
router
  policy-options
  begin
    community "NoExport" members "no-export"
  policy-statement "RemNoExport"
    default-action accept
    community remove "NoExport"
  exit
  exit
  commit
exit
```

This is applied as an export policy, so that the No-Export community is removed from all intra-AD routes advertised as updates to internal peers. The **vpn-apply-export** command must be included to ensure that the export policy is applied to routes belonging to VPN address families; in this case, MVPN-IPv4 routes.

```
# on PE-6:
configure
router
  bgp
    vpn-apply-export
    export "RemNoExport"
  exit
```

This policy should also be configured and applied on PE-1, so that intra-AD routes can be exported from MVPN PEs in AS 64496 to AS 64497.

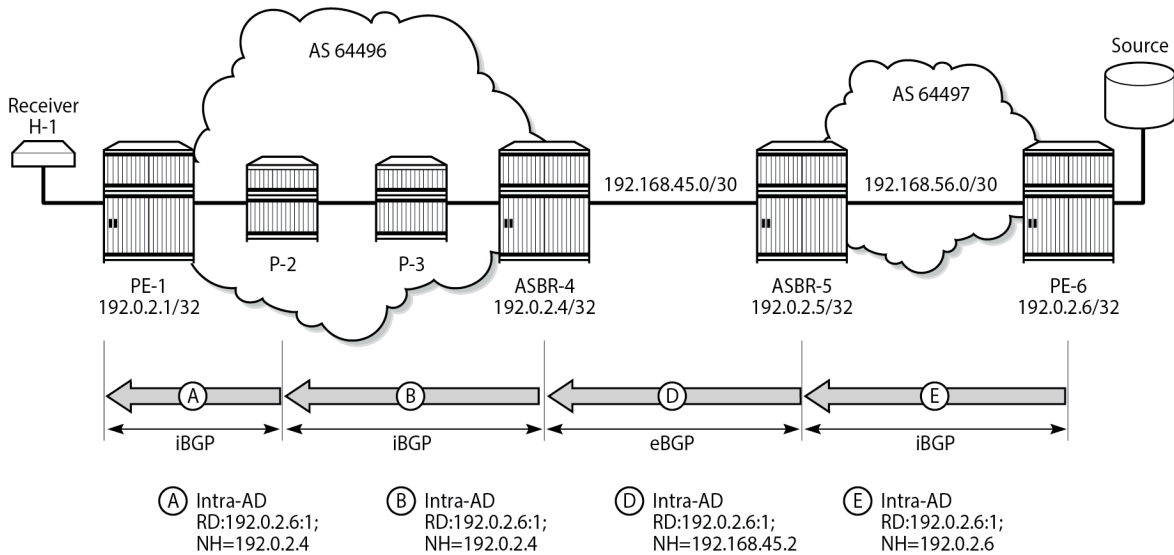
## Verification

### BGP MVPN Intra-AD Route Propagation

[Figure 43: BGP MVPN Intra-AD Route Advertisement](#) shows the propagation of the BGP MVPN intra-AD route from PE-6 to PE-1 across the AS boundary. The original route has the No-Export community removed on PE-6 because the export route policy is applied. ASBR-5 receives the route and forwards it to ASBR-4. The BGP next-hop attribute is changed at the AS boundary to the interface address of ASBR-5:

192.168.45.2. ASBR-4 forwards the intra-AD route to the RR on P-2, and changes the BGP next-hop attribute to its system address: 192.0.2.4. P-2 reflects the route to PE-1.

Figure 43: BGP MVPN Intra-AD Route Advertisement



PE-1 receives the route, and imports the route into VPRN 1 as the route target extended community matches the community configured in the MVPN context of the VPRN. PE-1 now uses the PTA contained within the intra-AD route to instantiate the provider tunnel.

The following output shows details of the MVPN intra-AD route received by PE-1, generated by PE-6.

```
# on PE-1:
*A:PE-1# show router bgp routes mvpn-ipv4 type intra-ad originator-ip 192.0.2.6 hunt
=====
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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP MVPN-IPv4 Routes
=====
RIB In Entries
-----
Route Type      : Intra-Ad
Route Dist.     : 192.0.2.6:1
Originator IP   : 192.0.2.6
Nexthop        : 192.0.2.4
Path Id         : None
From           : 192.0.2.2
Res. Nexthop    : 0.0.0.0
Local Pref.    : 100
Aggregator AS  : None
Atomic Aggr.   : Not Atomic
AIGP Metric     : None
Interface Name  : NotAvailable
Aggregator     : None
MED            : None
IGP Cost       : n/a
```

```

Connector      : None
Community      : target:64496:1
Cluster        : 192.0.2.2
Originator Id  : 192.0.2.4          Peer Router Id : 192.0.2.2
Origin         : IGP
Flags          : Used Valid Best
Route Source   : Internal
AS-Path        : 64497
Route Tag      : 0
Neighbor-AS    : 64497
DB Orig Val    : N/A              Final Orig Val : N/A
Source Class   : 0                Dest Class     : 0
Add Paths Send : Default
Last Modified  : 00h03m08s
VPRN Imported  : 1
-----
PMSI Tunnel Attributes :
Tunnel-type    : LDP P2MP LSP
Flags          : Type: RNVE(0) BM: 0 U: 0 Leaf: not required
MPLS Label     : 0
Root-Node      : 192.0.2.6        LSP-ID         : 8193
-----
RIB Out Entries
-----
Routes : 1
=====

```

## P2MP LDP LSP Signaling

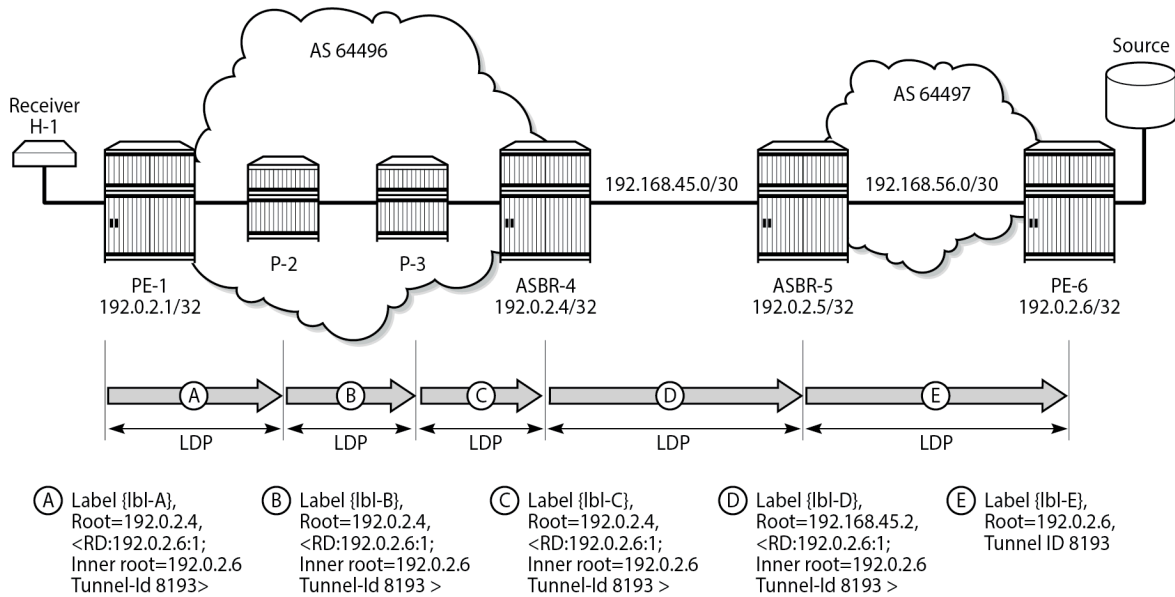
The PTA lists the tunnel type as a P2MP LDP LSP. A P2MP label mapping message is originated on PE-1, with LSP-ID 8193, and the root of the mLDP tree is PE-6: 192.0.2.6. However, PE-1 does not have a route to PE-6, because inter-AS model B VPNs do not require the system addresses of the PEs to be advertised into the neighboring AS.

The intra-AD MVPN route is used to determine the path of the label mapping message from PE-1 toward PE-6. This is comparable to the unicast routing case, where a VPN-IPv4 labeled route is used to determine the path to the source.

The BGP next-hop of the intra-AD route is the system address of ASBR-4, so this can be used as the root address of the mLDP LSP, and the actual root can be contained inside the label mapping message as an inner root. The inner root becomes an opaque value that is known to the originator and receiver of the label mapping message.

[Figure 44: P2MP LDP Label Mapping](#) shows the path taken by the label mapping message from PE-1 to PE-6.

Figure 44: P2MP LDP Label Mapping



27575b

P-2 and P-3 do not have either a unicast or multicast (intra-AD) route toward PE-6, but have a route to the outer root ASBR-4. The label mapping message is forwarded from PE-1 to ASBR-4 via P-2 and P-3. At each hop, a label is allocated and a label binding entry is created. The labels can be different on PE-1, P-2, P-3, ASBR-4, ASBR-5, and PE-6. In this case, the labels are: {lbl-A}=524282, {lbl-B}=524283, {lbl-C}=524283, {lbl-D}=524279, and {lbl-E}=524279. In the following sections, the debug outputs are achieved using the following debug command:

```
debug
  router "Base"
    ldp
      peer <peer-ip-address>
        packet
          label detail
        exit
      exit
    exit
  exit
exit
exit
```

where <peer-ip-address> is the system address of the LDP peer.

### LDP Hop PE-1 to P-2

The following output shows a debug of the P2MP label mapping message sent from PE-1 to P-2 upon receipt of the BGP MVPN intra-AD route.

```
# on PE-1:
9 2024/01/16 16:35:14.504 UTC MINOR: DEBUG #2001 Base LDP
"LDP: LDP
Send Label Mapping packet (msgId 86) to 192.0.2.2:0
```



```
Protocol version = 1
Label 524282 advertised for the following FECs
P2MP: root = 192.0.2.4, T: 8, L: 25 (RD: 0x1c00002060001, InnerRoot: 192.0.2.6 T: 1, L: 4,
TunnelId: 8193)
"
```

The advertised label is 524282: the ingress label on PE-1. The P2MP root address is that of the BGP next-hop of the intra-AD route, that is, the ASBR-4 system address. T: 8 signifies that the FEC type is 8, VPN-recursive FEC, and L: 25 is the length of the opaque value. The opaque value contains the route distinguisher (RD) of the intra-AD route plus inner root 192.0.2.6 and a second opaque value: a type 1 (T:1) generic of length L = 4 bytes, containing the tunnel ID 8193.

The format of the type 8 opaque value aligns with the representation in [Table 2: mLDP Message Opaque Value Types in MVPN Model B](#):

*<ASBR-4, Opaque type 8 <RD, PE-6 Opaque type 1 <Tunnel-ID>>>*.

The LDP binding table of PE-1 is shown in the following output:

```
# on PE-1:
*A:PE-1# show router ldp bindings active p2mp ipv4 opaque-type vpn-recursive

=====
LDP Bindings (IPv4 LSR ID 192.0.2.1)
(IPv6 LSR ID ::)
=====
Label Status:
  U - Label In Use, N - Label Not In Use, W - Label Withdrawn
  WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
  e - Label ELC
FEC Flags:
  LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
  BA - ASBR Backup FEC
=====
VPN Recursive with Generic IPv4 P2MP Bindings (Active)
=====
P2MP-Id          RD
InnerRootAddr    Interface
RootAddr         Op
IngLbl           EgrLbl
EgrNH            EgrIf/LspId
-----
8193             192.0.2.6:1
192.0.2.6       73728
192.0.2.4       Pop
524282          --
--              --
-----
No. of VPN Recursive with Generic IPv4 P2MP Active Bindings: 1
=====
```

This shows the VPN-recursive FEC binding with both root address of ASBR-4 and inner root of PE-6.

## LDP Hop P-2 to P-3

On P-2, the label mapping messages received from PE-1 and advertised toward P-3 are shown in the following output.

```
# on P-2:
```

```

16 2024/01/16 16:35:15.745 UTC MINOR: DEBUG #2001 Base LDP
"LDP: LDP
Recv Label Mapping packet (msgId 86) from 192.0.2.1:0
Protocol version = 1
Label 524282 advertised for the following FECs
P2MP: root = 192.0.2.4, T: 8, L: 25 (RD: 0x1c00002060001, InnerRoot: 192.0.2.6 T: 1, L: 4,
TunnelId: 8193)
"

17 2024/01/16 16:35:15.745 UTC MINOR: DEBUG #2001 Base LDP
"LDP: LDP
Send Label Mapping packet (msgId 84) to 192.0.2.3:0
Protocol version = 1
Label 524283 advertised for the following FECs
P2MP: root = 192.0.2.4, T: 8, L: 25 (RD: 0x1c00002060001, InnerRoot: 192.0.2.6 T: 1, L: 4,
TunnelId: 8193)
"

```

The received message matches the advertised label from PE-1, and the label mapping message toward P-3 (192.0.2.3) is again a VPN-recursive FEC type. P-3 does not have a route to PE-6, but has a route to ASBR-4.

The following output shows the LDP label mapping for the VPN-recursive FEC on P-2.

```

# on P-2:
*A:P-2# show router ldp bindings active p2mp ipv4 opaque-type vpn-recursive

=====
LDP Bindings (IPv4 LSR ID 192.0.2.2)
(IPv6 LSR ID ::)
=====
Label Status:
  U - Label In Use, N - Label Not In Use, W - Label Withdrawn
  WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
  e - Label ELC
FEC Flags:
  LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
  BA - ASBR Backup FEC
=====
VPN Recursive with Generic IPv4 P2MP Bindings (Active)
=====
P2MP-Id          RD
InnerRootAddr    Interface
RootAddr         Op
IngLbl           EgrLbl
EgrNH            EgrIf/LspId
-----
8193             192.0.2.6:1
192.0.2.6        Unknw
192.0.2.4        Swap
524283           524282
192.168.12.1     1/1/c2/1
-----
No. of VPN Recursive with Generic IPv4 P2MP Active Bindings: 1
=====

```

The following debug messages show the received and transmitted LDP label mapping message on P-3. The received label matches the advertised label from the previous debug output for P-2.

```

# on P-3:

```

```

1 2024/01/16 16:35:14.764 UTC MINOR: DEBUG #2001 Base LDP
"LDP: LDP
Recv Label Mapping packet (msgId 84) from 192.0.2.2:0
Protocol version = 1
Label 524283 advertised for the following FECs
P2MP: root = 192.0.2.4, T: 8, L: 25 (RD: 0x1c00002060001, InnerRoot: 192.0.2.6 T: 1, L: 4,
TunnelId: 8193)
"

2 2024/01/16 16:35:14.764 UTC MINOR: DEBUG #2001 Base LDP
"LDP: LDP
Send Label Mapping packet (msgId 80) to 192.0.2.4:0
Protocol version = 1
Label 524283 advertised for the following FECs
P2MP: root = 192.0.2.4, T: 8, L: 25 (RD: 0x1c00002060001, InnerRoot: 192.0.2.6 T: 1, L: 4,
TunnelId: 8193)
"

```

Again, the VPN-recursive FEC on P-3 is shown in the following output:

```

# on P-3:
*A:P-3# show router ldp bindings active p2mp ipv4 opaque-type vpn-recursive

=====
LDP Bindings (IPv4 LSR ID 192.0.2.3)
(IPv6 LSR ID ::)
=====
Label Status:
  U - Label In Use, N - Label Not In Use, W - Label Withdrawn
  WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
  e - Label ELC
FEC Flags:
  LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
  BA - ASBR Backup FEC
=====
VPN Recursive with Generic IPv4 P2MP Bindings (Active)
=====
P2MP-Id          RD
InnerRootAddr    Interface
RootAddr         Op
IngLbl           EgrLbl
EgrNH            EgrIf/LspId
-----
8193             192.0.2.6:1
192.0.2.6       Unknw
192.0.2.4       Swap
524283          524283
192.168.23.1   1/1/c2/1
-----
No. of VPN Recursive with Generic IPv4 P2MP Active Bindings: 1
=====

```

## ASBR-4

ASBR-4 is the root of the mLDP tree within AS 64496. Upon receipt of an mLDP label mapping message containing this FEC element, ASBR-4 recognizes that it is the root and that the opaque value is a VPN-recursive opaque value. ASBR-4 parses the VPN-recursive opaque value and extracts the root value: PE-6 plus the RD.

ASBR-4 does not have a unicast route to PE-6, so it must use the multicast VPN intra-AD route. This route contains an NLRI that has the IP address of PE-6, along with the BGP next-hop. As there may be multiple valid MVPN intra-ADs held by ASBR-4, the RD extracted from the mLDP label mapping message is used as a match to identify the MVPN intra-AD route held in the ASBR FIB.

ASBR-4 creates an mLDP mapping message containing a VPN-recursive FEC whose opaque value has the inner root address of PE-6, and a root address of ASBR-5.

The following output shows the label mapping messages at ASBR-4.

```
# on ASBR-4:
12 2024/01/16 16:35:14.922 UTC MINOR: DEBUG #2001 Base LDP
"LDP: LDP
Recv Label Mapping packet (msgId 80) from 192.0.2.3:0
Protocol version = 1
Label 524283 advertised for the following FECs
P2MP: root = 192.0.2.4, T: 8, L: 25 (RD: 0x1c00002060001, InnerRoot: 192.0.2.6 T: 1, L: 4,
TunnelId: 8193)
"

13 2024/01/16 16:35:14.923 UTC MINOR: DEBUG #2001 Base LDP
"LDP: Binding
Sending Label mapping label 524279 for P2MP: root = 192.168.45.2, T: 8, L: 25 (RD:
0x1c00002060001, InnerRoot: 192.0.2.6 T: 1, L: 4, TunnelId: 8193)
to peer 192.0.2.5:0."

14 2024/01/16 16:35:14.923 UTC MINOR: DEBUG #2001 Base LDP
"LDP: LDP
Send Label Mapping packet (msgId 78) to 192.0.2.5:0
Protocol version = 1
Label 524279 advertised for the following FECs
P2MP: root = 192.168.45.2, T: 8, L: 25 (RD: 0x1c00002060001, InnerRoot: 192.0.2.6 T: 1, L: 4,
TunnelId: 8193)
"
```

The label mapping message uses a format of the opaque type listed in [Table 2: mLDP Message Opaque Value Types in MVPN Model B](#), where the new root is now ASBR-5, and the inner root address remains the PE-6 system address:

*<ASBR-5, Opaque type 8 <RD, PE-6 Opaque type 1 <Tunnel-ID>>>*

At ASBR-4, the root changes from ASBR-4 to ASBR-5. ASBR-4 essentially becomes a leaf node with root at ASBR-5.

The following output shows the label binding output at ASBR-4.

```
# on ASBR-4:
*A:ASBR-4# show router ldp bindings active p2mp ipv4 opaque-type vpn-recursive

=====
LDP Bindings (IPv4 LSR ID 192.0.2.4)
(IPv6 LSR ID ::)
=====

Label Status:
  U - Label In Use, N - Label Not In Use, W - Label Withdrawn
  WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
  e - Label ELC
FEC Flags:
  LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
  BA - ASBR Backup FEC
```

```

=====
VPN Recursive with Generic IPv4 P2MP Bindings (Active)
=====
P2MP-Id                RD
InnerRootAddr          Interface
RootAddr               Op
IngLbl                 EgrLbl
EgrNH                  EgrIf/LspId
-----
8193                    192.0.2.6:1
192.0.2.6              Unknw
192.0.2.4 (LF)        Push
--                     524283
192.168.34.1          1/1/c2/1

8193                    192.0.2.6:1
192.0.2.6              Unknw
192.168.45.2 (UF)    Swap
524279                 Stitched
--                     --
-----
No. of VPN Recursive with Generic IPv4 P2MP Active Bindings: 2
=====

```

The label binding message received from the downstream router P-3 is known as the Lower FEC (LF). The label binding message forwarded to ASBR-5 has allocated a label and is stored as the Upper FEC (UF).

To create a non-segmented mLDP LSP, a label swap action must occur at ASBR-4, where the leaf of the P2MP LSP that has a root at ASBR-5 must be stitched to the P2MP LSP that has a root at ASBR-4 and leaf on PE-1. To achieve this, the UF label is swapped to the LF label. They are stitched using the common RD. If there are multiple lower FECs for the same RD at the ASBR, then ASBR-4 acts as a replication point. This stitching action can be seen in the EgrLbl field of the UF entry.

## ASBR-5

ASBR-5 receives the label mapping message from ASBR-4. This contains a label mapping plus the opaque value with a VPN-recursive FEC type 8. The root address is a local address, so the recursive FEC is parsed and the root address of PE-6 is extracted.

```

# on ASBR-5:
12 2024/01/16 16:35:13.506 UTC MINOR: DEBUG #2001 Base LDP
"LDP: LDP
Recv Label Mapping packet (msgId 78) from 192.0.2.4:0
Protocol version = 1
Label 524279 advertised for the following FECs
P2MP: root = 192.168.45.2, T: 8, L: 25 (RD: 0x1c00002060001, InnerRoot: 192.0.2.6 T: 1, L: 4,
TunnelId: 8193)
"

```

ASBR-5 has a route to the PE-6 address 192.0.2.6 in the forwarding table.

ASBR-5 therefore constructs an mLDP label mapping message with FEC element containing the address of PE-6 as the root address. This is seen in the following output, where the opaque type is type 1. The opaque value is the tunnel ID contained in the original intra-AD MVPN route.

```

# on ASBR-5:
13 2024/01/16 16:35:13.506 UTC MINOR: DEBUG #2001 Base LDP
"LDP: LDP

```

```
Send Label Mapping packet (msgId 72) to 192.0.2.6:0
Protocol version = 1
Label 524279 advertised for the following FECs
P2MP: root = 192.0.2.6, T: 1, L: 4, TunnelId: 8193
"
```

This compares to the representation for opaque type 1 from [Table 2: mLDP Message Opaque Value Types in MVPN Model B](#):

<PE-6 Opaque type 1 <Tunnel-ID>>.

The following output taken from ASBR-5 shows the stitching of the recursive label mapping received from ASBR-4 to the generic IPv4 label mapping sent to PE-6. The LF label received from ASBR-4 (524279) is stitched to the UF label (524279) via the common root address of 192.0.2.6.

```
# on ASBR-5:
*A:ASBR-5# show router ldp bindings active p2mp ipv4

=====
LDP Bindings (IPv4 LSR ID 192.0.2.5)
(IPv6 LSR ID ::)
=====
Label Status:
  U - Label In Use, N - Label Not In Use, W - Label Withdrawn
  WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
  e - Label ELC
FEC Flags:
  LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
  BA - ASBR Backup FEC
=====
LDP Generic IPv4 P2MP Bindings (Active)
=====
P2MP-Id          Interface
RootAddr         Op
IngLbl           EgrLbl
EgrNH            EgrIf/LspId
-----
8193              Unknw
192.0.2.6 (UF)   Swap
524279           Stitched
--              --
-----
No. of Generic IPv4 P2MP Active Bindings: 1
=====
---snip---
=====
VPN Recursive with Generic IPv4 P2MP Bindings (Active)
=====
P2MP-Id          RD
InnerRootAddr    Interface
RootAddr         Op
IngLbl           EgrLbl
EgrNH            EgrIf/LspId
-----
8193              192.0.2.6:1
192.0.2.6        Unknw
192.168.45.2 (LF) Push
--              524279
192.168.45.1     1/1/c2/1
-----
No. of VPN Recursive with Generic IPv4 P2MP Active Bindings: 1
```

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

## LDP Hop ASBR-5 to PE-6

For completeness, the following debug output on PE-6 shows the receipt of the mLDP label mapping message from ASBR-5, which contains the system address of PE-6 as the root address.

```
# on PE-6:  
8 2024/01/16 16:35:14.819 UTC MINOR: DEBUG #2001 Base LDP  
"LDP: LDP  
Recv Label Mapping packet (msgId 72) from 192.0.2.5:0  
Protocol version = 1  
Label 524279 advertised for the following FECs  
P2MP: root = 192.0.2.6, T: 1, L: 4, TunnelId: 8193  
"
```

The label binding output on PE-6 shows that the operation is a push operation. This is expected because PE-6 is the root node of the P2MP LSP.

```
# on PE-6:  
*A:PE-6# show router ldp bindings active p2mp ipv4 opaque-type generic
```

```
=====
LDP Bindings (IPv4 LSR ID 192.0.2.6)
(IPv6 LSR ID ::)
=====
```

Label Status:

```
U - Label In Use, N - Label Not In Use, W - Label Withdrawn
WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
e - Label ELC
```

FEC Flags:

```
LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
BA - ASBR Backup FEC
```

```
=====
LDP Generic IPv4 P2MP Bindings (Active)
=====
```

P2MP-Id	Interface
RootAddr	Op
IngLbl	EgrLbl
EgrNH	EgrIf/LspId
-----	-----
8193	73728
192.0.2.6	Push
--	524279
192.168.56.1	1/1/c2/1
-----	-----

```
No. of Generic IPv4 P2MP Active Bindings: 1
=====
```

## PIM status

Traffic is forwarded into multicast group 239.255.0.1 from the source using address 172.16.67.2.

An IGMPv3 group membership report is generated by the receiver H-1 and is shown on PE-1.

```
# on PE-1:
*A:PE-1# show router 1 igmp group
=====
IGMP Interface Groups
=====
(172.16.67.2,239.255.0.1)                UpTime: 0d 00:03:20
  Fwd List  : int-PE-1-VPRN-1-H-1
-----
Entries : 1
=====
---snip---
```

The status of the PIM group for VPRN 1 for group 239.255.0.1 is shown in the following output.

```
# on PE-1:
*A:PE-1# show router 1 pim group
=====
Legend:  A = Active   S = Standby
=====
PIM Groups ipv4
=====
Group Address          Type          Spt Bit  Inc Intf      No.0ifs
  Source Address        RP           State    Inc Intf(S)
-----
239.255.0.1            (S,G)                mpls-if-73728  1
  172.16.67.2
-----
Groups : 1
=====
```

The incoming interface is an MPLS interface: mpls-if-73728. This is a PIM tunnel interface, as shown in the following output:

```
# on PE-1:
*A:PE-1# show router 1 pim tunnel-interface
=====
PIM Interfaces ipv4
=====
Interface              Originator Address  Adm  Opr  Transport Type
-----
mpls-if-73728           192.0.2.6           Up   Up   Rx-IPMSI
mpls-virt-if-1005857   192.0.2.1           Up   Up   Tx-IPMSI
-----
Interfaces : 2
=====
```

The originator address is 192.0.2.6, which is the root address of the mLDP tunnel on PE-6-the non-segmented mLDP tunnel.

For completeness, the PIM status of the group 239.255.0.1 on PE-6 is as follows:

```
# on PE-6:
*A:PE-6# show router 1 pim group
=====
Legend:  A = Active   S = Standby
```



```
=====
PIM Groups ipv4
=====
Group Address          Type          Spt Bit   Inc Intf   No.0ifs
Source Address        RP
-----
239.255.0.1           (S,G)                int-PE-6-VPRN* 1
172.16.67.2
-----
Groups : 1
=====
* indicates that the corresponding row element may have been truncated.
```

## Conclusion

Inter-AS multicast within a VPRN can be achieved using non-segmented mLDP trees. This chapter provides the configuration for inter-AS model B MVPN. The example also shows the associated commands, debug, and outputs, which can be used for verifying and troubleshooting.

# NG-MVPN Inter-AS Model C Using Non-Segmented mLDP Tunnels

This chapter provides information about NG-MVPN Inter-AS Model C Using Non-Segmented mLDP Tunnels.

Topics in this chapter include:

- [Applicability](#)
- [Overview](#)
- [Configuration](#)
- [Conclusion](#)

## Applicability

The information and configuration in this chapter are based on SR OS Release 23.10.R2.

No specific configuration is required to support non-segmented multicast Label Distribution Protocol (mLDP) for inter-AS model C. However, recursive-opaque mLDP Forwarding Equivalence Class (FEC) functionality must be supported.

The configuration of multicast in a VPRN is described in the [NG-MVPN Configuration with MPLS](#) chapter.

## Overview

Multicast in an inter-AS model C network can be implemented using Rosen, where the set of Multicast Distribution Trees (MDTs) are signaled using Protocol Independent Multicast Source-Specific Mode (PIM-SSM).

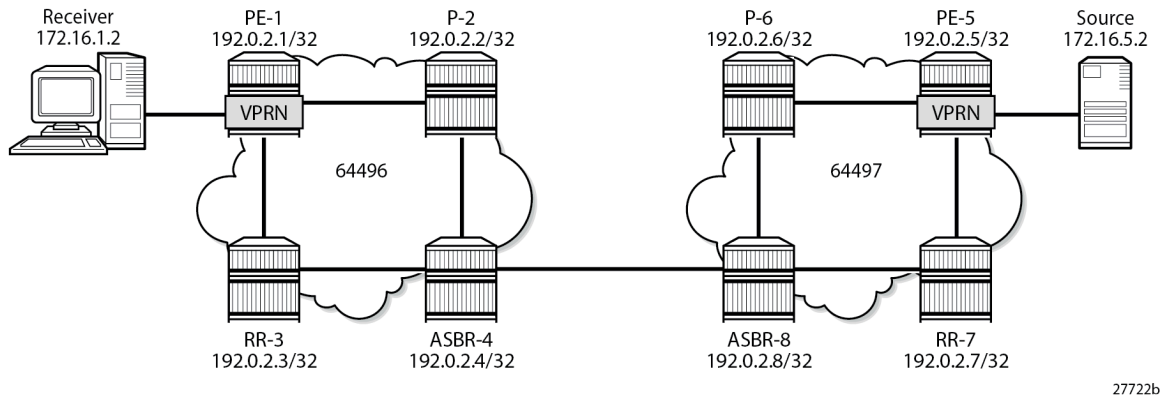
It is also possible to create MDTs between PEs in different Autonomous Systems (ASs) for Next-Generation MVPN (NG-MVPN) using non-segmented dynamic multicast LDP (mLDP) trees, where the root and leaf PEs are in different ASs. This chapter describes the configuration of MVPN services between PEs in different ASs using NG-MVPN techniques.

[Figure 45: NG-MVPN Inter-AS Model C](#) shows an example of a network comprising routers in two neighboring ASs, modeled as an inter-AS model C network; see the [Inter-AS VPRN Model C](#) chapter.

A VPRN instance exists in PE-1 and PE-5, with a single source and receiver connected. RR-3 is an off the data path Route Reflector (RR) for the vpn-ipv4 and mvpn-ipv4 address families, and ASBR-4 is RR for the label-ipv4 address family in AS 64496. PE-1 is client to both RRs. The situation in AS 64497 is similar.

The multicast source located at address 172.16.5.2 in AS 64497 generates multicast traffic with multicast group address 239.255.0.1, and the receiver at address 172.16.1.2 in AS 64496 becomes a member of the multicast group using IGMPv3 signaling. No multicast protocols are configured on the ASBRs.

Figure 45: NG-MVPN Inter-AS Model C



In an MVPN, a transport tunnel is signaled that carries multicast traffic from source PE to any receiver PE that has multicast routers attached, or hosts that want to become multicast group members. In this case, an mLDP tunnel is signaled between source PE and destination PEs to carry multicast traffic across the multi-AS provider network.

An mLDP Provider Multicast Service Interface (PMSI), also known as the provider tunnel, is established between PEs that declare membership to the MVPN, by generating and advertising an MVPN type 1 intra-AD (Auto-Discovery) BGP route. This route contains a PMSI Tunnel Attribute (PTA) that provides the tunnel type, the root node, and the LSP ID. Upon receipt of the intra-AD route, the receiving PE checks the route validity so that it can be imported into the VPRN instance. If the route is valid, the receiving router signals a point-to-multipoint (P2MP) LDP label mapping message toward the root address contained within the PTA of the intra-AD route. At the root, MPLS-encapsulated multicast traffic is forwarded to the downstream router by pushing on a label received from the downstream router.

Inter-AS model C unicast requires the RRs in peer ASes to establish a multi-hop EBGP session over which vpn-ipv4 routes can then be exchanged. If RRs are not used, PEs in the peer ASes can exchange the vpn-ipv4 and mvpn-ipv4 routes over a multi-hop eBGP session. Therefore, model C also requires the PE system addresses to be advertised across the AS boundary first.

The path for inter-AS unicast traffic is resolved using vpn-ipv4 routes learnt via the multi-hop eBGP session between RR-3 and RR-7 and label-ipv4 routes learnt via ASBR-4 and ASBR-8. In a unicast environment, traffic from PE-1 to PE-5 takes the following path:

- The tunnel from PE-1 to ASBR-4 comprises of three MPLS labels: the MPLS transport label to reach ASBR-4, followed by the MPLS label to reach PE-5 (bgp label-ipv4 learnt from ASBR-4), followed by the label associated with the vpn-ipv4 route.
- ASBR-4 pops the outer-most MPLS transport label and forwards to ASBR-8.
- ASBR-8 swaps the outer MPLS label (bgp label-ipv4) with the MPLS transport label to reach PE-5.
- PE-5 pops the outer MPLS label and uses the VPRN label to forward traffic to the appropriate VPRN.

Unlike unicast, multicast requires a non-segmented provider tunnel for traffic to be routed from the root PE toward the receiver routers. This means that the tunnel must traverse multiple ASes without decapsulating and encapsulating labels at the AS boundaries, and therefore, must be non-segmented.

If the provider tunnel uses mLDP, the receivers initiate the signaling by sending an LDP label mapping message along the control path toward the root. This follows the path of the intra-AD route that advertises the root of the ingress PMSI (I-PMSI). An mLDP label mapping message consists of an allocated label L,

with FEC element <X,Y>, where X identifies the root node and Y is the opaque value, so the P2MP label mapping can be denoted as <X,Y,L>.

The FEC element contains the root address of the LSP plus a variable-length opaque value. The opaque value contains information meaningful to the root and leaf routers, but not to intermediate routers; for example, a P2MP LSP ID or a nested opaque value.

The root address is the system address of PE-5 (the router that advertised the intra-AD route) which is reachable via ASBR-4.

However, in an mLDP environment, each router must take part in the signaling of the P2MP LSP, but not every router has an MVPN route, so any mLDP label mapping message received by P-2 to the root address of PE-5 is dropped. A solution to this is described in RFC 6512, *Using Multipoint LDP When the Backbone Has No Route to the Root*, as follows.

PE-1 signals an mLDP LSP as if the root is at ASBR-4. P-2 has a route to ASBR-4 because it participates in the IGP through the same IGP instance. The actual root address located in PE-5 is encapsulated in the mLDP label mapping message originated by PE-1 as an inner root address. This is a recursive FEC type 7; the root (PE-5) is encapsulated within a FEC element as an opaque value.

ASBR-4 receives the intra-AD MVPN route advertised by PE-5 with BGP next-hop ASBR-8.

[Table 3: mLDP Message Opaque Value Types in MVPN inter-AS Model C](#) shows the opaque value types used in MVPN inter-AS model C.

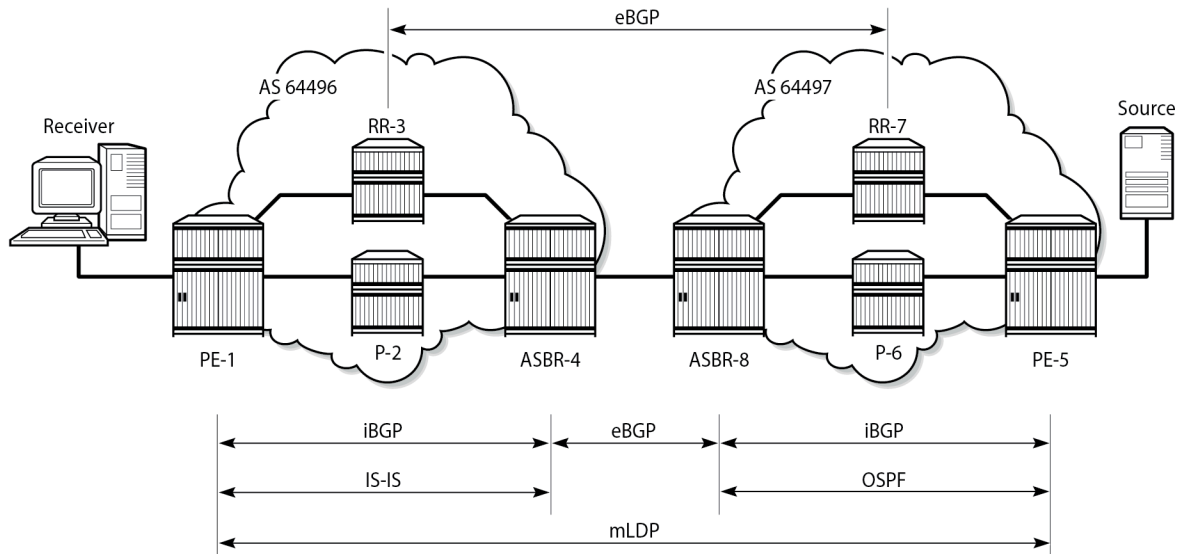
*Table 3: mLDP Message Opaque Value Types in MVPN inter-AS Model C*

Opaque Type	Opaque Name	Use	FEC Element Representation
1	Generic	VPRN local AS	Root, Opaque<P2MP ID>
7	Recursive FEC	Inter-AS model C mLDP	<ASBR, Opaque <Root, P2MP ID>>

## Configuration

[Figure 46: Inter-AS MVPN Protocol Requirements](#) shows the required protocol configuration and peering.

Figure 46: Inter-AS MVPN Protocol Requirements

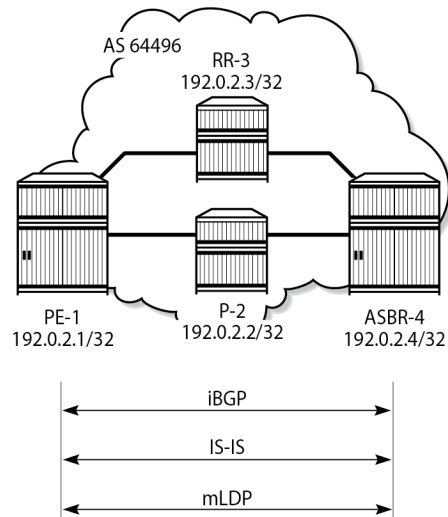


27723b

## AS 64496

Figure 47: AS 64496 Protocols shows the protocol requirements for AS 64496.

Figure 47: AS 64496 Protocols



27724b

## Router Interface and IS-IS Configuration

The first step is to configure the router interfaces and IS-IS as the Interior Gateway Protocol (IGP) in AS 64496.

The router interfaces for PE-1 are configured as follows:

```
# on PE-1
configure
  router "Base"
    interface "int-PE-1-P-2"
      address 192.168.12.1/30
      port 1/1/c1/1
    exit
    interface "int-PE-1-RR-3"
      address 192.168.13.1/30
      port 1/1/c2/1
    exit
    interface "system"
      address 192.0.2.1/32
    exit
  exit
```

Each interface is configured to run IS-IS as the IGP, as follows. Each router is configured as a level 2 router.

```
# on PE-1
configure
  router "Base"
    isis 0
      level-capability level-2
      area-id 49.0001
      advertise-passive-only
      level 2
        wide-metrics-only
      exit
    interface "system"
      passive
    exit
    interface "int-PE-1-P-2"
      interface-type point-to-point
    exit
    interface "int-PE-1-RR-3"
      interface-type point-to-point
    exit
  no shutdown
  exit
```

The configuration for all other nodes in the AS is the same, other than the IP addresses. The IP addresses can be derived from Figure 1.

## LDP Configuration

LDP is used as the MPLS protocol and must be enabled on each router interface, except for the VPN RR and the interfaces to the RR, as follows:

```
# on PE-1
configure
```

```

router "Base"
  ldp
    import-pmsi-routes
    exit
    tcp-session-parameters
    exit
    interface-parameters
      interface "int-PE-1-P-2" dual-stack
        ipv4
          fec-type-capability
          p2mp-ipv4 enable
          exit
          no shutdown
        exit
      no shutdown
    exit
  exit
  targeted-session
  exit
  no shutdown
exit

```

This configuration must be repeated on each of the other routers in the AS. Because LDP is used as the provider tunnel interface for multicast traffic, the previously indicated interfaces must also support P2MP LDP tunnels. Therefore, the FEC type capability for IPv4 P2MP tunnels must be enabled. The default value is enable, but is included in the preceding output for clarity.

## BGP Configuration

### RR-3 Configuration

RR-3 is the route reflector for the vpn-ipv4 and mvpn-ipv4 address families within AS 64496, and peers with PE-1, but not with ASBR-4 and P-2. The cluster ID is set to ensure that RR-3 is an RR. For supporting inter-AS VPRN model C, RR-3 also peers with RR-7 in AS 64497 for the same address families, through a multi-hop eBGP session.

```

# on RR-3
configure
  router "Base"
    autonomous-system 64496
    bgp
      loop-detect discard-route
      rapid-withdrawal
      split-horizon
      rapid-update vpn-ipv4 mvpn-ipv4
      group "EBGP-vpn-mvpn"
        local-address 192.0.2.3
        neighbor 192.0.2.7
          family vpn-ipv4 mvpn-ipv4
          multihop 10
          peer-as 64497
        exit
      exit
    group "IBGP-vpn-mvpn"
      cluster 192.0.2.3
      neighbor 192.0.2.1
        family vpn-ipv4 mvpn-ipv4
        peer-as 64496

```

```

        exit
    exit
exit
exit

```

## ASBR-4 Configuration

Although ASBR-4 is the AS border router for AS 64496 (so that it peers with ASBR-8 in AS 64497), ASBR-4 is also configured as RR for the label-ipv4 address family within AS 64496, and peers with PE-1, as follows:

```

# on ASBR-4
configure
  router "Base"
    autonomous-system 64496
    bgp
      loop-detect discard-route
      rapid-withdrawal
      split-horizon
      rapid-update vpn-ipv4 mvpn-ipv4
      rib-management
        label-ipv4
          route-table-import "pol-imp-PE-and-RR-pfxs"
        exit
      exit
    group "EBGP-label"
      neighbor 192.168.48.2
        family label-ipv4
        peer-as 64497
        advertise-inactive
        export "pol-exp-PE-and-RR-pfxs"
      exit
    exit
  group "IBGP-label"
    cluster 192.0.2.4
    neighbor 192.0.2.1
      family label-ipv4
      peer-as 64496
    exit
  exit
exit
exit
exit

```

The policies ensure that the system addresses of the PEs and the RRs are exchanged between AS 64496 and AS 64497. In this example, ASBR-4 generates the labeled routes but pops the label associated with RR-3; see the *Pop-Label for /32 Label-IPv4 BGP routes* chapter in the *7450 ESS, 7750 SR, and 7950 XRS Unicast Routing Protocols Advanced Configuration Guide for Classic CLI*.

```

# on ASBR-4
configure
  router "Base"
    policy-options
      begin
      prefix-list "PE-pfxs"
        prefix 192.0.2.1/32 exact
        prefix 192.0.2.4/32 exact
      exit
      prefix-list "RR-pfxs"
        prefix 192.0.2.3/32 exact

```



```

exit
policy-statement "pol-exp-PE-and-RR-pfxs"
  entry 10
    from
      prefix-list "PE-pfxs" "RR-pfxs"
    exit
    action accept
  exit
exit
policy-statement "pol-imp-PE-and-RR-pfxs"
  entry 20
    from
      prefix-list "RR-pfxs"
    exit
    action accept
    advertise-label pop
  exit
  entry 30
    from
      prefix-list "PE-pfxs"
    exit
    action accept
  exit
exit
commit
exit

```

## PE-1 configuration

PE-1 is a BGP client of RR-3 and ASBR-4, as follows:

```

# on PE-1
configure
  router "Base"
    autonomous-system 64496
    bgp
      loop-detect discard-route
      rapid-withdrawal
      split-horizon
      rapid-update vpn-ipv4 mvpn-ipv4
      group "IBGP-label"
        neighbor 192.0.2.4
          family label-ipv4
          peer-as 64496
      exit
    group "IBGP-vpn-mvpn"
      neighbor 192.0.2.3
        family vpn-ipv4 mvpn-ipv4
        peer-as 64496
    exit
  exit
exit
exit

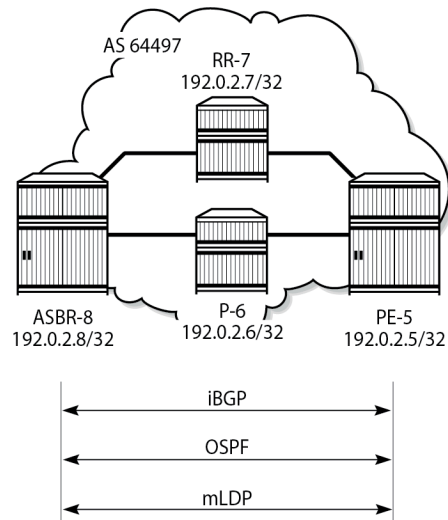
```

P-2 does not have any BGP configuration; it participates in the MPLS data plane for switching unicast and multicast customer traffic, and does not have any services defined.

## AS 64497

Figure 48: AS 64497 Protocols shows the protocol requirements for AS 64497.

Figure 48: AS 64497 Protocols



27725b

## Router Interface and OSPF Configuration

The first step is to configure router interfaces and OSPF on each router shown in [Figure 48: AS 64497 Protocols](#). All router interfaces are members of a single backbone area: area 0.0.0.0.

The following router interfaces are configured on PE-5:

```
# on PE-5
configure
  router "Base"
    interface "int-PE-5-P-6"
      address 192.168.56.1/30
      port 1/1/c2/1
    exit
    interface "int-PE-5-RR-7"
      address 192.168.57.1/30
      port 1/1/c1/1
    exit
    interface "system"
      address 192.0.2.5/32
    exit
  exit
```

On PE-5, OSPF is configured as follows:

```
# on PE-5
configure
  router "Base"
    ospf 0
```

```

    area 0.0.0.0
      interface "int-PE-5-P-6"
        interface-type point-to-point
      exit
      interface "int-PE-5-RR-7"
        interface-type point-to-point
      exit
      interface "system"
      exit
    exit
  no shutdown
exit

```

## LDP Configuration

LDP is used as the MPLS protocol and must be enabled on each router interface, except on the RR and the interfaces toward the RR, as follows:

```

# on PE-5
configure
  router "Base"
    ldp
      import-pmi-routes
      exit
      tcp-session-parameters
      exit
      interface-parameters
        interface "int-PE-5-P-6" dual-stack
          ipv4
            fec-type-capability
            p2mp-ipv4 enable
          exit
          no shutdown
        exit
      exit
    exit
  targeted-session
  exit
  no shutdown
exit

```

This configuration must be repeated on each of the other routers in the AS, and the FEC type capability for IPv4 P2MP tunnels must again be enabled. The default value is enable, but is included in the preceding output for clarity.

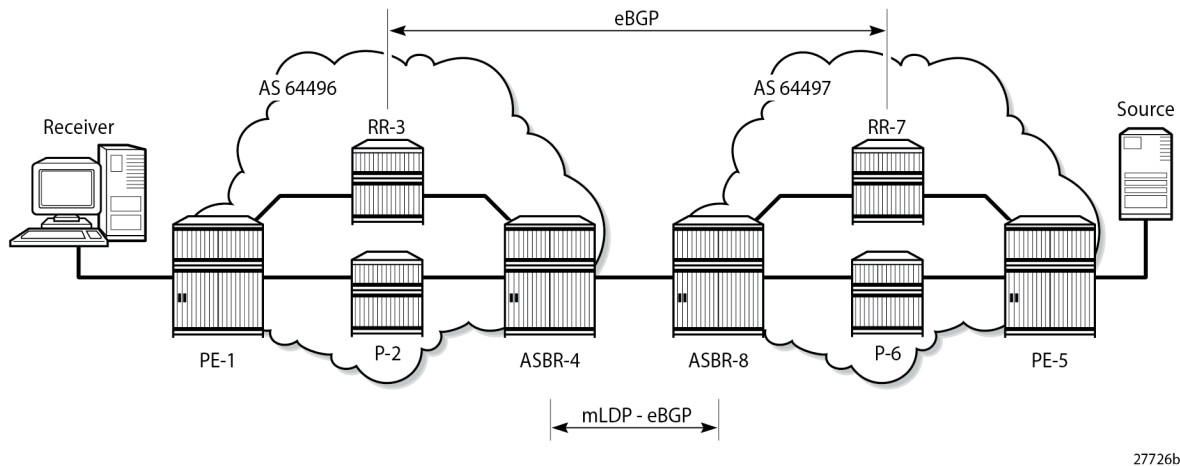
## BGP Configuration

The BGP configuration for AS 64497 is mirrored from AS 64496, so the configuration is similar, meaning that RR-7 is RR for the vpn-ipv4 and mvpn-ipv4 address families, and ASBR-8 is RR for the label-ipv4 address family, with PE-5 being the client. P-6 is not a BGP client for RR-7 and ASBR-8.

## Inter-AS Configuration

Figure 49: Inter-AS Protocols shows the protocols required between ASBR-4 and ASBR-8. The interface addresses are used for the LDP transport and the BGP speaker peer addresses.

Figure 49: Inter-AS Protocols



27726b

## LDP Peering

LDP is configured as the MPLS protocol between ASBR-4 and ASBR-8. On ASBR-4, the interface toward ASBR-8 has LDP enabled, as follows. The interface address should be used as the LDP transport address because the system address of ASBR-8 is not known to ASBR-4.

```
# on ASBR-4
configure
  router "Base"
    ldp
      import-pmsi-routes
      exit
      tcp-session-parameters
      exit
      interface-parameters
        interface "int-ASBR-4-ASBR-8" dual-stack
          ipv4
            fec-type-capability
            p2mp-ipv4 enable
            exit
            transport-address interface
            no shutdown
          exit
        exit
      exit
    exit
  targeted-session
  exit
  no shutdown
exit
```

The P2MP FEC type capability for P2MP LDP is shown. This is the default value.

For completeness, the LDP configuration on ASBR-8 for the interface toward ASBR-4 is as follows:

```
# on ASBR-8
configure
  router "Base"
    ldp
      import-pmsi-routes
      exit
      tcp-session-parameters
      exit
      interface-parameters
        interface "int-ASBR-8-ASBR-4" dual-stack
          ipv4
            fec-type-capability
              p2mp-ipv4 enable
            exit
            transport-address interface
              no shutdown
          exit
        no shutdown
      exit
    exit
  targeted-session
  exit
no shutdown
exit
```

The LDP session is successfully established at ASBR-4, as follows:

```
*A:ASBR-4# show router ldp session 192.0.2.8

=====
LDP IPv4 Sessions
=====
Peer LDP Id      Adj Type  State      Msg Sent  Msg Recv  Up Time
-----
192.0.2.8:0     Link     Established  181       182       0d 00:07:35
-----
No. of IPv4 Sessions: 1
=====

=====
LDP IPv6 Sessions
=====
Peer LDP Id      Adj Type  State      Msg Sent  Msg Recv  Up Time
-----
No Matching Entries Found
=====
```

For completeness, the LDP session from ASBR-8 toward ASBR-4 is shown in the following output:

```
*A:ASBR-8# show router ldp session 192.0.2.4

=====
LDP IPv4 Sessions
=====
Peer LDP Id      Adj Type  State      Msg Sent  Msg Recv  Up Time
-----
192.0.2.4:0     Link     Established  184       185       0d 00:07:45
-----
```

```

-----
No. of IPv4 Sessions: 1
=====
LDP IPv6 Sessions
=====
Peer LDP Id      State      Msg Sent    Msg Recv    Up Time
Adj Type
-----
No Matching Entries Found
=====

```

When a label mapping message is received for an LDP FEC prefix, the next-hop for a FEC prefix is resolved in the routing table. The FEC is installed in the Label Information Base (LIB) if the next-hop matches a /32 route table entry.

The local interface configuration results in a route being installed with a subnet mask matching the interface configuration. In this case, the ASBR-to-ASBR route is 192.168.48.0/30.

For LDP to resolve the LDP FEC egress next-hop on ASBR-4, a /32 route matching the egress next-hop address is required in the FIB. For that purpose, on ASBR-4, a static route is configured for the /32 address located on ASBR-8, as follows:

```

# on ASBR-4
configure
  router "Base"
    static-route-entry 192.168.48.2/32
      next-hop 192.168.48.2
      no shutdown
  exit

```

Similarly, a static route on ASBR-8 is configured for the /32 address located on ASBR-4, as follows:

```

# on ASBR-8
configure
  router "Base"
    static-route-entry 192.168.48.1/32
      next-hop 192.168.48.1
      no shutdown
  exit

```

The following output shows that the static route is installed in the ASBR-4 Routing Information Base (RIB):

```

*A:ASBR-4# show router route-table protocol static
=====
Route Table (Router: Base)
=====
Dest Prefix[Flags]      Type  Proto  Age      Pref
Next Hop[Interface Name] Metric
-----
192.168.48.2/32        Remote Static  00h00m00s  5
192.168.48.2          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

```

## eBGP Peering - ASBR-4 to ASBR-8

The following shows the eBGP peering configuration for ASBR-4. The peer address is the interface address of ASBR-8.

```
# on ASBR-4
configure
  router "Base"
    autonomous-system 64496
    bgp
      group "EBGP-label"
        neighbor 192.168.48.2
          family label-ipv4
          peer-as 64497
          advertise-inactive
          export "pol-exp-PE-and-RR-pfxs"
        exit
      exit
    exit
  exit
```

Similarly, the BGP configuration for ASBR-8 peering toward ASBR-4 is as follows:

```
# on ASBR-8
configure
  router "Base"
    autonomous-system 64497
    bgp
      group "EBGP-label"
        neighbor 192.168.48.1
          family label-ipv4
          peer-as 64496
          advertise-inactive
          export "pol-exp-PE-and-RR-pfxs"
        exit
      exit
    exit
  exit
```

The BGP peering session between ASBR-4 and ASBR-8 is verified as follows. In this example, only label-ipv4 routes are exchanged.

```
*A:ASBR-4# show router bgp summary group "EBGP-label"
=====
BGP Router ID:192.0.2.4      AS:64496      Local AS:64496
=====
BGP Admin State      : Up      BGP Oper State      : Up
Total Peers          : 1
Total IPv4 Remote Rts : 0      Total IPv4 Rem. Active Rts : 0
Total IPv6 Remote Rts : 0      Total IPv6 Rem. Active Rts : 0
Total IPv4 Backup Rts : 0      Total IPv6 Backup Rts   : 0
Total LblIpv4 Rem Rts : 3      Total LblIpv4 Rem. Act Rts : 3
Total LblIpv6 Rem Rts : 0      Total LblIpv6 Rem. Act Rts : 0
Total LblIpv4 Bkp Rts : 0      Total LblIpv6 Bkp Rts   : 0
Total VPN-IPv4 Rem. Rts : 0      Total VPN-IPv4 Rem. Act. Rts: 0
```

```

Total VPN-IPv6 Rem. Rts : 0          Total VPN-IPv6 Rem. Act. Rts: 0
Total VPN-IPv4 Bkup Rts : 0          Total VPN-IPv6 Bkup Rts      : 0

Total MVPN-IPv4 Rem Rts : 0          Total MVPN-IPv4 Rem Act Rts : 0
Total MVPN-IPv6 Rem Rts : 0          Total MVPN-IPv6 Rem Act Rts : 0
Total MDT-SAFI Rem Rts  : 0          Total MDT-SAFI Rem Act Rts  : 0
Total McIPv4 Remote Rts : 0          Total McIPv4 Rem. Active Rts: 0
Total McIPv6 Remote Rts : 0          Total McIPv6 Rem. Active Rts: 0
Total McVpnIPv4 Rem Rts : 0          Total McVpnIPv4 Rem Act Rts : 0
Total McVpnIPv6 Rem Rts : 0          Total McVpnIPv6 Rem Act Rts : 0

Total EVPN Rem Rts      : 0          Total EVPN Rem Act Rts     : 0
Total L2-VPN Rem. Rts   : 0          Total L2VPN Rem. Act. Rts   : 0
Total MSPW Rem Rts      : 0          Total MSPW Rem Act Rts     : 0
Total RouteTgt Rem Rts : 0          Total RouteTgt Rem Act Rts : 0
Total FlowIpv4 Rem Rts  : 0          Total FlowIpv4 Rem Act Rts  : 0
Total FlowIpv6 Rem Rts  : 0          Total FlowIpv6 Rem Act Rts  : 0
Total FlowVpvn4 Rem Rts : 0          Total FlowVpvn4 Rem Act Rts : 0
Total FlowVpvn6 Rem Rts : 0          Total FlowVpvn6 Rem Act Rts : 0
Total Link State Rem Rts: 0          Total Link State Rem Act Rts: 0
Total SrPlcyIpv4 Rem Rts: 0          Total SrPlcyIpv4 Rem Act Rts: 0
Total SrPlcyIpv6 Rem Rts: 0          Total SrPlcyIpv6 Rem Act Rts: 0

```

=====

BGP Summary

=====

Legend : D - Dynamic Neighbor

=====

Neighbor  
Description

	AS	PktRcvd	InQ	Up/Down	State	Rcv/Act/Sent	(Addr Family)
			OutQ				
-----							
192.168.48.2	64497	13	0	00h02m12s	3/3/3	(LbL-IPv4)	
		13	0				
-----							

## eBGP Peering - RR-3 to RR-7

The following shows the configuration required for establishing an eBGP multi-hop peering session from RR-3 to RR-7. The peer address is the system address of RR-7.

```

# on RR-3
configure
router "Base"
static-route-entry 192.0.2.7/32
indirect 192.0.2.4
tunnel-next-hop
resolution disabled
exit
no shutdown
exit
exit
autonomous-system 64496
bgp
loop-detect discard-route
rapid-withdrawal
split-horizon
rapid-update vpn-ipv4 mvpn-ipv4
group "EBGP-vpn-mvpn"

```



```

        local-address 192.0.2.3
        neighbor 192.0.2.7
            family vpn-ipv4 mvpn-ipv4
            multihop 10
            peer-as 64497
        exit
    exit
    ---snip---
    no shutdown
exit
exit

```

Similarly, the BGP configuration for RR-7 multi-hop peering with RR-3 is as follows:

```

# on RR-7
configure
router "Base"
    static-route-entry 192.0.2.3/32
        indirect 192.0.2.8
        tunnel-next-hop
        resolution disabled
    exit
    no shutdown
exit
autonomous-system 64497
bgp
    loop-detect discard-route
    rapid-withdrawal
    split-horizon
    rapid-update vpn-ipv4 mvpn-ipv4
    group "EBGP-vpn-mvpn"
        local-address 192.0.2.7
        neighbor 192.0.2.3
            family vpn-ipv4 mvpn-ipv4
            multihop 10
            peer-as 64496
        exit
    exit
    ---snip---
    no shutdown
exit
exit

```

This peering session is verified as follows. Only vpn-ipv4 and mvpn-ipv4 routes are exchanged.

```

*A:RR-3# show router bgp summary group "EBGP-vpn-mvpn"
=====
BGP Router ID:192.0.2.3      AS:64496      Local AS:64496
=====
BGP Admin State      : Up      BGP Oper State      : Up
Total Peers          : 1
Total IPv4 Remote Rts : 0      Total IPv4 Rem. Active Rts : 0
Total IPv6 Remote Rts : 0      Total IPv6 Rem. Active Rts : 0
Total IPv4 Backup Rts : 0      Total IPv6 Backup Rts      : 0
Total LblIpv4 Rem Rts : 0      Total LblIpv4 Rem. Act Rts : 0
Total LblIpv6 Rem Rts : 0      Total LblIpv6 Rem. Act Rts : 0
Total LblIpv4 Bkp Rts : 0      Total LblIpv6 Bkp Rts      : 0

Total VPN-IPv4 Rem. Rts : 1      Total VPN-IPv4 Rem. Act. Rts: 0
Total VPN-IPv6 Rem. Rts : 0      Total VPN-IPv6 Rem. Act. Rts: 0
Total VPN-IPv4 Bkup Rts : 0      Total VPN-IPv6 Bkup Rts    : 0

```

```

Total MVPN-IPv4 Rem Rts : 1          Total MVPN-IPv4 Rem Act Rts : 0
Total MVPN-IPv6 Rem Rts : 0          Total MVPN-IPv6 Rem Act Rts : 0
Total MDT-SAFI Rem Rts : 0           Total MDT-SAFI Rem Act Rts : 0
Total McIPv4 Remote Rts : 0          Total McIPv4 Rem. Active Rts: 0
Total McIPv6 Remote Rts : 0          Total McIPv6 Rem. Active Rts: 0
Total McVpnIPv4 Rem Rts : 0          Total McVpnIPv4 Rem Act Rts : 0
Total McVpnIPv6 Rem Rts : 0          Total McVpnIPv6 Rem Act Rts : 0

Total EVPN Rem Rts      : 0           Total EVPN Rem Act Rts      : 0
Total L2-VPN Rem. Rts   : 0           Total L2VPN Rem. Act. Rts   : 0
Total MSPW Rem Rts     : 0           Total MSPW Rem Act Rts     : 0
Total RouteTgt Rem Rts : 0           Total RouteTgt Rem Act Rts : 0
Total FlowIpv4 Rem Rts : 0           Total FlowIpv4 Rem Act Rts : 0
Total FlowIpv6 Rem Rts : 0           Total FlowIpv6 Rem Act Rts : 0
Total FlowVpvn4 Rem Rts : 0          Total FlowVpvn4 Rem Act Rts : 0
Total FlowVpvn6 Rem Rts : 0          Total FlowVpvn6 Rem Act Rts : 0
Total Link State Rem Rts: 0           Total Link State Rem Act Rts: 0
Total SrPlcyIpv4 Rem Rts: 0          Total SrPlcyIpv4 Rem Act Rts: 0
Total SrPlcyIpv6 Rem Rts: 0          Total SrPlcyIpv6 Rem Act Rts: 0

```

=====  
BGP Summary  
=====

Legend : D - Dynamic Neighbor  
=====

Neighbor  
Description

Neighbor Description	AS	PktRcvd PktSent	InQ OutQ	Up/Down	State	Rcv/Act/Sent (Addr Family)
----- 192.0.2.7	64497	33 34	0 0	00h13m55s	1/0/1	(VpnIPv4)
-----					1/0/2	(MvpnIPv4)
-----						

## VPRN Configuration

The VPRN service configuration for PE-1 and PE-5 is as follows:

```

# on PE-1
configure
  service
    vprn 1 name "VPRN 1" customer 1 create
    interface "int-PE-1-VPRN-1-H-1" create
      address 172.16.1.1/30
      sap 1/1/c3/1 create
    exit
  exit
  bgp-ipvpn
    mpls
      auto-bind-tunnel
      resolution-filter
      ldp
    exit
    resolution filter
  exit
  route-distinguisher 64496:1
  vrf-target export target:64496:1 import target:64497:1
  no shutdown
exit

```

```

exit
igmp
  interface "int-PE-1-VPRN-1-H-1"
    version 3 # default
    ssm-translate
      grp-range 239.255.0.0 239.255.255.255
      source 172.16.5.2
    exit
  exit
  static
    group 239.255.0.1
    starg
  exit
  exit
  no shutdown
exit
no shutdown
exit
pim
  rp
    static
    exit
    bsr-candidate
      shutdown
    exit
    rp-candidate
      shutdown
    exit
  exit
  no shutdown
exit
mvpn
  auto-discovery default
  c-mcast-signaling bgp
  no intersite-shared
  mdt-type receiver-only
  provider-tunnel
    inclusive
    mldp
    no shutdown
  exit
  exit
  vrf-target unicast
  exit
exit
no shutdown
exit

```

```

# on PE-5
configure
  service
    vprn 1 name "VPRN 1" customer 1 create
    interface "int-PE-5-VPRN-1-S-5" create
      address 172.16.5.1/30
      sap 1/1/c3/1 create
    exit
  exit
  bgp-ipvpn
    mpls
      auto-bind-tunnel
      resolution-filter
      ldp

```

```

        exit
        resolution filter
    exit
    route-distinguisher 64497:1
    vrf-target export target:64497:1 import target:64496:1
    no shutdown
    exit
exit
igmp
    interface "int-PE-5-VPRN-1-S-5"
        version 3 # default
        no shutdown
    exit
    no shutdown
exit
pim
    interface "int-PE-5-VPRN-1-S-5"
    exit
    rp
        static
        exit
        bsr-candidate
        shutdown
        exit
        rp-candidate
        shutdown
        exit
    exit
    no shutdown
exit
mvpn
    auto-discovery default
    c-mcast-signaling bgp
    no intersite-shared
    mdt-type sender-only
    provider-tunnel
        inclusive
        mldp
        no shutdown
    exit
    exit
    vrf-target unicast
    exit
exit
no shutdown
exit

```

## Route Policy for MVPN Routes

The use of non-segmented LDP provider tunnels requires that intra-AD routes must be advertised across the AS boundary between PEs. Each intra-AD route generated by a PE that is a member of an MVPN contains the well-known community "No-Export", which prevents a BGP speaker from advertising the route across an AS boundary to another external BGP speaker.

In inter-AS model C, the RRs must support the MVPN address family. If the RR receives an intra-AD route containing the No-Export community, this route is not advertised to any external peer. A route policy is required to remove the No-Export community before the route can be advertised across the AS boundary to a BGP speaker that has negotiated the MVPN address family capability.

In the following example, a *RemNoExport* policy is defined, which removes the No-Export community, using the community remove action as a default action:

```
# on PE-1 and PE-5
configure
  router "Base"
    policy-options
      begin
        community "NoExport"
          members "no-export"
        exit
        policy-statement "RemNoExport"
          default-action accept
            community remove "NoExport"
          exit
        exit
      commit
    exit
```

This policy is applied as an export policy, so that the No-Export community is removed from all intra-AD routes advertised as updates to internal peers. The **vpn-apply-export** command must be included to ensure that the export policy is applied to routes belonging to VPN address families; in this case, mvpn-ipv4 routes.

```
# on PE-5
configure
  router "Base"
    bgp
      group "IBGP-vpn-mvpn"
        neighbor 192.0.2.7
          family vpn-ipv4 mvpn-ipv4
          peer-as 64497
          vpn-apply-export
          export "RemNoExport"
        exit
      exit
    no shutdown
  exit
```

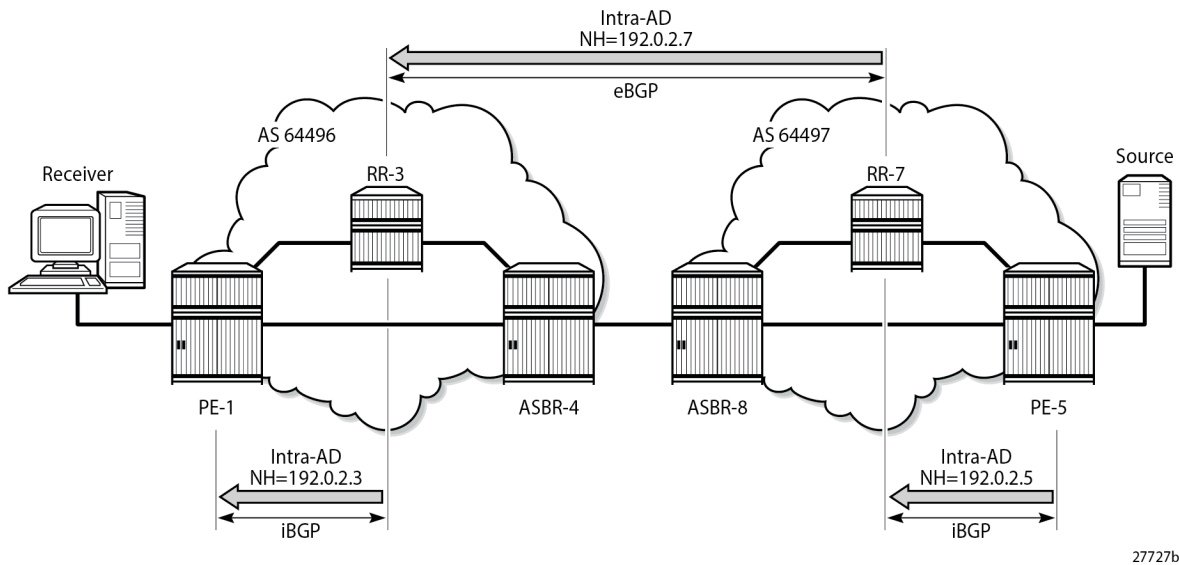
This policy should also be configured and applied on PE-1, so that intra-AD routes can be exported from MVPN PEs in AS 64496 to AS 64497.

## Verification

### BGP MVPN Intra-AD Route Propagation

[Figure 50: BGP MVPN Intra-AD Route Advertisement](#) shows the propagation of the BGP MVPN intra-AD route from PE-5 to PE-1 across the AS boundary. The original route has the No-Export community removed on PE-5, because the export route policy is applied. RR-3 receives the route via RR-7 and reflects it to PE-1. The BGP next-hop attribute is changed by RR-7 to its system address: 192.0.2.7. RR-3 reflects the intra-AD route to PE-1, also changing the BGP next-hop attribute to its system address: 192.0.2.3.

Figure 50: BGP MVPN Intra-AD Route Advertisement



27727b

PE-1 receives the route, and imports the route into VPRN 1 because the route target extended community matches the community configured in the MVPN context of the VPRN. PE-1 now uses the PTA contained within the intra-AD route to instantiate the provider tunnel.

The following output shows details of the MVPN intra-AD route received by PE-1, generated by PE-5:

```
*A:PE-1# show router bgp routes mvpn-ipv4 type intra-ad originator-ip 192.0.2.5 hunt
=====
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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP MVPN-IPv4 Routes
=====
RIB In Entries
-----
Route Type      : Intra-Ad
Route Dist.     : 64497:1
Originator IP   : 192.0.2.5
Nexthop         : 192.0.2.3
Path Id         : None
From            : 192.0.2.3
Res. Nexthop    : 0.0.0.0
Local Pref.     : 100
Aggregator AS   : None
Atomic Aggr.    : Not Atomic
AIGP Metric     : None
Connector       : None
Community       : target:64497:1
Cluster         : No Cluster Members
Originator Id   : None
Origin          : IGP
Interface Name  : NotAvailable
Aggregator      : None
MED             : None
IGP Cost        : n/a
Peer Router Id  : 192.0.2.3
```

```

Flags          : Used Valid Best
Route Source   : Internal
AS-Path        : 64497
Route Tag      : 0
Neighbor-AS    : 64497
DB Orig Val    : N/A                Final Orig Val : N/A
Source Class   : 0                  Dest Class     : 0
Add Paths Send : Default
Last Modified  : 00h05m55s
VPRN Imported  : 1
-----
PMSI Tunnel Attributes :
Tunnel-type      : LDP P2MP LSP
Flags           : Type: RNVE(0) BM: 0 U: 0 Leaf: not required
MPLS Label      : 0
Root-Node       : 192.0.2.5          LSP-ID        : 8193
-----
RIB Out Entries
-----
Routes : 1
=====

```

## P2MP LDP LSP Signaling

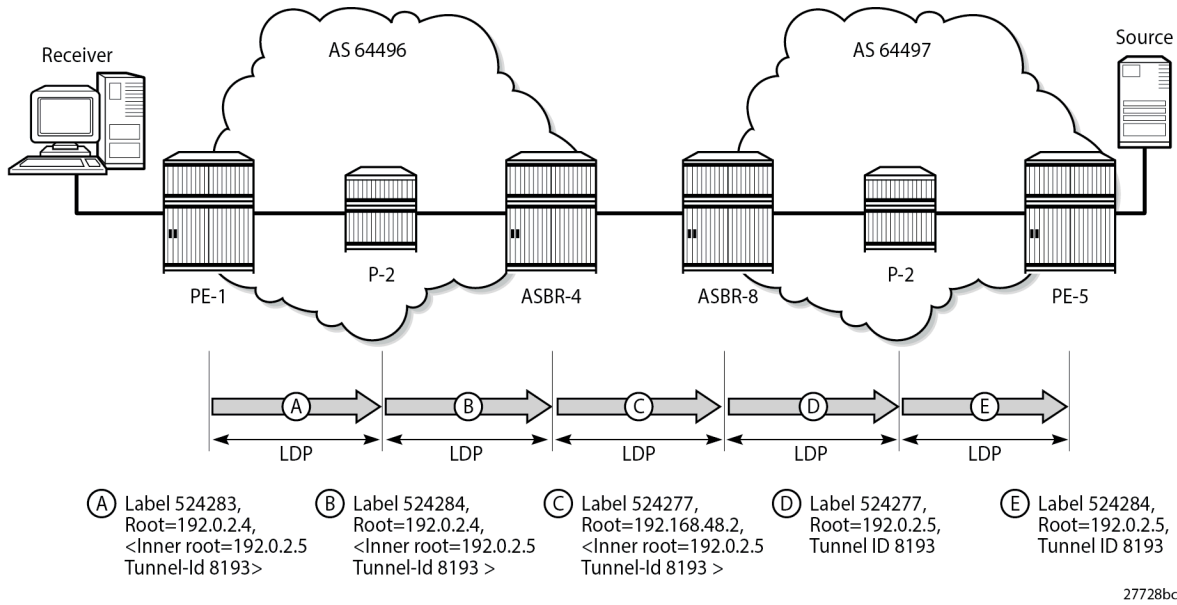
The PTA lists the tunnel type as an LDP P2MP LSP. A P2MP label mapping message is originated on PE-1, with LSP ID 8193, with inner-root PE-5 and outer-root ASBR-4.

The intra-AD MVPN route is used to determine the path of the label mapping message from PE-1 toward PE-5. This is similar to the unicast routing case, where a vpn-ipv4 labeled route is used to determine the path to the source.

The BGP next-hop of the intra-AD route is the system address of ASBR-4, so this can be used as the root address of the mLDP LSP, and the actual root can be contained inside the label mapping message as an inner root. The inner root becomes an opaque value that is known to the originator and receiver of the label mapping message. PE-1 thus generates a recursive-opaque type=7 FEC: <ASBR-4, <PE-5, P2MP-id>>.

[Figure 51: P2MP LDP Label Mapping](#) shows the path taken by the label mapping message from PE-1 to PE-5.

Figure 51: P2MP LDP Label Mapping



P-2 does not have a BGP-LBL route to reach PE-5, but has an IGP route to reach the outer root ASBR-4. The label mapping message is forwarded from PE-1 to ASBR-4 via P-2. At each hop, a label is allocated and a label binding entry is created. In the following sections, the debug outputs are achieved using the following debug command:

```
debug
router "Base"
  ldp
    peer <peer-ip-address>
    packet
    label detail
  exit
exit
exit
exit
exit
exit
```

where <peer-ip-address> is the system address of the LDP peer.

### LDP Hop PE-1 to P-2

The following output shows a debug of the P2MP label mapping message sent from PE-1 to P-2 upon receipt of the BGP MVPN intra-AD route:

```
7 2024/02/21 00:07:41.996 UTC MINOR: DEBUG #2001 Base LDP
"LDP: LDP
Send Label Mapping packet (msgId 233) to 192.0.2.2:0
Protocol version = 1
Label 524283 advertised for the following FECs
P2MP: root = 192.0.2.4, T: 7, L: 17 (InnerRoot: 192.0.2.5 T: 1, L: 4, TunnelId: 8193)
"
```



The advertised label is 524283: the ingress label on PE-1. The P2MP root address is the next hop for the BGP-LBL route to reach PE-5 (192.0.2.5), namely ASBR-4 (192.0.2.4). T: 7 signifies that the FEC type is 7, GRT-recursive FEC, and L: 17 is the length of the opaque value. The opaque value contains the inner root 192.0.2.5 and a second opaque value: a type 1 (T: 1) generic FEC of length L = 4 bytes, containing the tunnel ID 8193.

The format of the type 7 opaque value aligns with the representation in Table 1:

<ASBR-4, Opaque type 7 <PE-5, Opaque type 1 <Tunnel-ID>>>.

The LDP binding table of PE-1 is shown in the following output:

```
*A:PE-1# show router ldp bindings active p2mp ipv4 opaque-type grt-recursive
=====
LDP Bindings (IPv4 LSR ID 192.0.2.1)
      (IPv6 LSR ID ::)
=====
Label Status:
  U - Label In Use, N - Label Not In Use, W - Label Withdrawn
  WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
  e - Label ELC
FEC Flags:
  LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
  BA - ASBR Backup FEC
=====
LDP GRT Recursive with Generic IPv4 P2MP Bindings (Active)
=====
P2MP-Id
InnerRootAddr      Interface
RootAddr           0p
IngLbl             EgrLbl
EgrNH              EgrIf/LspId
-----
8193
192.0.2.5          73728
192.0.2.4          Pop
524283             --
--                 --
-----
No. of GRT Recursive with Generic IPv4 P2MP Active Bindings: 1
=====
```

The preceding output shows the GRT-recursive FEC binding with both the root address of ASBR-4 and the inner root address of PE-5.

## LDP Hop P-2 to ASBR-4

At P-2, the label mapping messages received from PE-1 and advertised toward ASBR-4 are shown in the following output:

```
1 2024/02/21 00:07:41.135 UTC MINOR: DEBUG #2001 Base LDP
"LDP: LDP
Recv Label Mapping packet (msgId 233) from 192.0.2.1:0
Protocol version = 1
Label 524283 advertised for the following FECs
P2MP: root = 192.0.2.4, T: 7, L: 17 (InnerRoot: 192.0.2.5 T: 1, L: 4, TunnelId: 8193)
```

```
"
2 2024/02/21 00:07:41.135 UTC MINOR: DEBUG #2001 Base LDP
"LDP: LDP
Send Label Mapping packet (msgId 227) to 192.0.2.4:0
Protocol version = 1
Label 524284 advertised for the following FECs
P2MP: root = 192.0.2.4, T: 7, L: 17 (InnerRoot: 192.0.2.5 T: 1, L: 4, TunnelId: 8193)
"
```

The received message matches the advertised label from PE-1, and the label mapping message toward ASBR-4 (192.0.2.4) is again a GRT-recursive FEC type.

The following output shows the LDP label mapping for the GRT-recursive FEC on P-2:

```
*A:P-2# show router ldp bindings active p2mp ipv4 opaque-type grt-recursive

=====
LDP Bindings (IPv4 LSR ID 192.0.2.2)
              (IPv6 LSR ID ::)
=====
Label Status:
  U - Label In Use, N - Label Not In Use, W - Label Withdrawn
  WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
  e - Label ELC
FEC Flags:
  LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
  BA - ASBR Backup FEC
=====
LDP GRT Recursive with Generic IPv4 P2MP Bindings (Active)
=====
P2MP-Id
InnerRootAddr      Interface
RootAddr           Op
IngLbl             EgrLbl
EgrNH              EgrIf/LspId
-----
8193
192.0.2.5          Unknw
192.0.2.4          Swap
524284             524283
192.168.12.1      1/1/c2/1
-----
No. of GRT Recursive with Generic IPv4 P2MP Active Bindings: 1
=====
```

## LDP Hop ASBR-4 to ASBR-8

ASBR-4 is the root of the mLDP tree in AS 64496. Upon receipt of an mLDP label mapping message containing this FEC element, ASBR-4 recognizes that it is the root and that the opaque value is a GRT-recursive opaque value. ASBR-4 parses the GRT-recursive opaque value and extracts the root value: PE-5.

ASBR-4 checks RTM for a route to reach the inner-root (PE-5), which is a BGP-LBL route with next-hop ASBR-8.

ASBR-4 creates an mLDP mapping message containing a GRT-recursive FEC whose opaque value has the inner root address of PE-5, and a root address of ASBR-8.

The following output shows the label mapping messages at ASBR-4:

```
1 2024/02/21 00:07:42.375 UTC MINOR: DEBUG #2001 Base LDP
"LDP: LDP
Recv Label Mapping packet (msgId 227) from 192.0.2.2:0
Protocol version = 1
Label 524284 advertised for the following FECs
P2MP: root = 192.0.2.4, T: 7, L: 17 (InnerRoot: 192.0.2.5 T: 1, L: 4, TunnelId: 8193)
"
```

```
2 2024/02/21 00:07:42.375 UTC MINOR: DEBUG #2001 Base LDP
"LDP: Binding
Sending Label mapping label 524277 for P2MP: root = 192.168.48.2, T: 7, L: 17
(InnerRoot: 192.0.2.5 T: 1, L: 4, TunnelId: 8193)
to peer 192.0.2.8:0."
```

```
3 2024/02/21 00:07:42.375 UTC MINOR: DEBUG #2001 Base LDP
"LDP: LDP
Send Label Mapping packet (msgId 142) to 192.0.2.8:0
Protocol version = 1
Label 524277 advertised for the following FECs
P2MP: root = 192.168.48.2, T: 7, L: 17 (InnerRoot: 192.0.2.5 T: 1, L: 4, TunnelId: 8193)
"
```

The label mapping message uses a format of the opaque type listed in [Table 3: mLDP Message Opaque Value Types in MVPN inter-AS Model C](#), where the new root is now ASBR-8, and the inner root address remains the PE-5 system address:

*<ASBR-8, Opaque type 7 <PE-5, Opaque type 1 <Tunnel-ID>>>*

At ASBR-4, the root changes from ASBR-4 to ASBR-8. ASBR-4 essentially becomes a leaf node with root at ASBR-8.

The following output shows the label binding output at ASBR-4:

```
*A:ASBR-4# show router ldp bindings active p2mp ipv4 opaque-type grt-recursive

=====
LDP Bindings (IPv4 LSR ID 192.0.2.4)
(IPv6 LSR ID ::)
=====
Label Status:
  U - Label In Use, N - Label Not In Use, W - Label Withdrawn
  WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
  e - Label ELC
FEC Flags:
  LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
  BA - ASBR Backup FEC
=====
LDP GRT Recursive with Generic IPv4 P2MP Bindings (Active)
=====
P2MP-Id
InnerRootAddr          Interface
RootAddr                Op
IngLbl                  EgrLbl
EgrNH                   EgrIf/LspId
-----
8193
192.0.2.5                Unknw
192.0.2.4 (LF)          Push
--                       524284
```

```

192.168.24.1                                1/1/c2/1

8193
192.0.2.5                                Unknw
192.168.48.2 (UF)                        Swap
524277                                    Stitched
--                                         --

-----
No. of GRT Recursive with Generic IPv4 P2MP Active Bindings: 2
=====

```

The label binding message received from the downstream router P-2 is known as the Lower FEC (LF). The label binding message forwarded to ASBR-8 has allocated a label and is stored as the Upper FEC (UF).

To create a non-segmented mLDP LSP, a label swap action must occur at ASBR-4, where the leaf of the P2MP LSP that has a root at ASBR-8 must be stitched to the P2MP LSP that has a root at ASBR-4 and leaf on PE-1. To achieve this, the UF label is swapped with the LF label. This stitching action is shown in the EgrLbl field of the UF entry.

## LDP Hop ASBR-8 to P-6

ASBR-8 receives the label mapping message from ASBR-4. This contains a label mapping plus the opaque value with a GRT-recursive FEC type 7. The root address is a local address, so the recursive FEC is parsed and the root address of PE-5 is extracted.

```

1 2024/02/21 00:07:42.158 UTC MINOR: DEBUG #2001 Base LDP
"LDP: LDP
Recv Label Mapping packet (msgId 142) from 192.0.2.4:0
Protocol version = 1
Label 524277 advertised for the following FECs
P2MP: root = 192.168.48.2, T: 7, L: 17 (InnerRoot: 192.0.2.5 T: 1, L: 4, TunnelId: 8193)
"

```

ASBR-8 has an IGP route to the PE-5 address (192.0.2.5) in the forwarding table.

Therefore, ASBR-8 constructs an mLDP label mapping message with FEC element containing the address of PE-5 as the root address. This is shown in the following output, where the opaque type is type 1. The opaque value is the tunnel ID contained in the original intra-AD MVPN route, which was contained in the lower FEC received from ASBR-4.

```

2 2024/02/21 00:07:42.159 UTC MINOR: DEBUG #2001 Base LDP
"LDP: LDP
Send Label Mapping packet (msgId 214) to 192.0.2.6:0
Protocol version = 1
Label 524277 advertised for the following FECs
P2MP: root = 192.0.2.5, T: 1, L: 4, TunnelId: 8193
"

```

This aligns with the representation for generic FEC type 1 from Table 1:

*<PE-5 Opaque type 1 <Tunnel-ID>>*

The following output taken from ASBR-8 shows the stitching of the recursive label mapping received from ASBR-4 to the generic IPv4 label mapping sent to P-6. The LF label received from ASBR-4 (524277) is stitched to the UF label (524277) via the common root address of 192.0.2.5.

```
*A:ASBR-8# show router ldp bindings active p2mp ipv4
```

```

=====
LDP Bindings (IPv4 LSR ID 192.0.2.8)
(IPv6 LSR ID ::)
=====
Label Status:
  U - Label In Use, N - Label Not In Use, W - Label Withdrawn
  WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
  e - Label ELC
FEC Flags:
  LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
  BA - ASBR Backup FEC
=====
LDP Generic IPv4 P2MP Bindings (Active)
=====
P2MP-Id          Interface
RootAddr         Op
IngLbl           EgrLbl
EgrNH            EgrIf/LspId
-----
8193             Unknw
192.0.2.5 (UF)   Swap
524277           Stitched
--              --
-----
No. of Generic IPv4 P2MP Active Bindings: 1
=====
---snip---
=====
LDP GRT Recursive with Generic IPv4 P2MP Bindings (Active)
=====
P2MP-Id          Interface
InnerRootAddr    Op
RootAddr         EgrLbl
IngLbl           EgrIf/LspId
EgrNH            EgrIf/LspId
-----
8193             Unknw
192.0.2.5         Push
192.168.48.2 (LF) 524277
--              524277
192.168.48.1     1/1/c3/1
-----
No. of GRT Recursive with Generic IPv4 P2MP Active Bindings: 1
=====

```

## LDP Hop P-6 to PE-5

At P-6, the label mapping messages received from ASBR-8 and advertised toward PE-5 are shown in the following output:

```

1 2024/02/21 00:07:41.068 UTC MINOR: DEBUG #2001 Base LDP
"LDP: LDP
Recv Label Mapping packet (msgId 214) from 192.0.2.8:0
Protocol version = 1
Label 524277 advertised for the following FECs
P2MP: root = 192.0.2.5, T: 1, L: 4, TunnelId: 8193

```

```
"  
2 2024/02/21 00:07:41.068 UTC MINOR: DEBUG #2001 Base LDP  
"LDP: LDP  
Send Label Mapping packet (msgId 218) to 192.0.2.5:0  
Protocol version = 1  
Label 524284 advertised for the following FECs  
P2MP: root = 192.0.2.5, T: 1, L: 4, TunnelId: 8193  
"
```

The following output on P-6 shows the LDP label mapping for this FEC:

```
*A:P-6# show router ldp bindings active p2mp ipv4 opaque-type generic
```

```
=====
```

```
LDP Bindings (IPv4 LSR ID 192.0.2.6)  
(IPv6 LSR ID ::)
```

```
=====
```

Label Status:

U - Label In Use, N - Label Not In Use, W - Label Withdrawn  
WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route  
e - Label ELC

FEC Flags:

LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,  
BA - ASBR Backup FEC

```
=====
```

```
LDP Generic IPv4 P2MP Bindings (Active)
```

```
=====
```

P2MP-Id	Interface
RootAddr	Op
IngLbl	EgrLbl
EgrNH	EgrIf/LspId
-----	
8193	Unknw
192.0.2.5	Swap
524284	524277
192.168.68.2	1/1/c2/1

```
-----
```

```
No. of Generic IPv4 P2MP Active Bindings: 1
```

```
=====
```

## PE-5

Finally, the following debug output on PE-5 shows the receipt of the mLDP label mapping message sent by P-6, which contains the system address of PE-5 as the root address:

```
7 2024/02/21 00:07:41.364 UTC MINOR: DEBUG #2001 Base LDP  
"LDP: LDP  
Recv Label Mapping packet (msgId 218) from 192.0.2.6:0  
Protocol version = 1  
Label 524284 advertised for the following FECs  
P2MP: root = 192.0.2.5, T: 1, L: 4, TunnelId: 8193  
"
```

The label binding output on PE-5 shows that the operation is a push operation. This is expected because PE-5 is the root node of the P2MP LSP.

```
*A:PE-5# show router ldp bindings active p2mp ipv4 opaque-type generic
=====
LDP Bindings (IPv4 LSR ID 192.0.2.5)
(IPv6 LSR ID ::)
=====
Label Status:
  U - Label In Use, N - Label Not In Use, W - Label Withdrawn
  WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
  e - Label ELC
FEC Flags:
  LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
  BA - ASBR Backup FEC
=====
LDP Generic IPv4 P2MP Bindings (Active)
=====
P2MP-Id          Interface
RootAddr         Op
IngLbl           EgrLbl
EgrNH            EgrIf/LspId
-----
8193             73728
192.0.2.5       Push
--              524284
192.168.56.2    1/1/c2/1
-----
No. of Generic IPv4 P2MP Active Bindings: 1
=====
```

## PIM status

Traffic is forwarded into multicast group 239.255.0.1 from the source using address 172.16.5.2.

An IGMPv3 group membership report is generated by the receiver and is shown on PE-1:

```
*A:PE-1# show router 1 igmp group
=====
IGMP Interface Groups
=====
(*,239.255.0.1)                               UpTime: 0d 00:08:28
  Fwd List  : int-PE-1-VPRN-1-H-1
-----
Entries : 1
=====
IGMP Host Groups
=====
No Matching Entries
=====
IGMP SAP Groups
=====
No Matching Entries
=====
IGMP SLA Profile Instance Groups
=====
No Matching Entries
```

The status of the PIM group for VPRN 1 for multicast group 239.255.0.1 is shown in the following output:

```
*A:PE-1# show router 1 pim group detail
=====
PIM Source Group ipv4
=====
Group Address      : 239.255.0.1
Source Address     : 172.16.5.2
RP Address         : 0
Advt Router       : 192.0.2.5
Flags             :                               Type           : (S,G)
Mode              : sparse
MRIB Next Hop    : 192.0.2.5
MRIB Src Flags   : remote
Keepalive Timer  : Not Running
Up Time          : 0d 00:08:28      Resolved By       : rtable-u

Up JP State      : Joined           Up JP Expiry      : 0d 00:00:27
Up JP Rpt       : Not Joined StarG  Up JP Rpt Override : 0d 00:00:00

Register State   : No Info
Reg From Anycast RP: No

Rpf Neighbor     : 192.0.2.5
Incoming Intf    : mpls-if-73728
Outgoing Intf List : int-PE-1-VPRN-1-H-1

Curr Fwding Rate : 78.560 kbps
Forwarded Packets : 3905           Discarded Packets : 0
Forwarded Octets  : 3834710       RPF Mismatches    : 0
Spt threshold    : 0 kbps          ECMP opt threshold : 7
Admin bandwidth  : 1 kbps

-----
Groups : 1
=====
```

The incoming interface is an MPLS interface: mpls-if-73728. This is a PIM tunnel interface, as shown in the following output:

```
*A:PE-1# show router 1 pim tunnel-interface
=====
PIM Interfaces ipv4
=====
Interface                Originator Address  Adm  Opr  Transport Type
-----
mpls-if-73728            192.0.2.5          Up   Up   Rx-IPMSI
mpls-virt-if-1005857    192.0.2.1          Up   Up   Tx-IPMSI
-----
Interfaces : 2
=====
```

The originator address is 192.0.2.5, which is the root address of the mLDP tunnel on PE-5 — the non-segmented mLDP tunnel.

For completeness, the PIM status of the multicast group 239.255.0.1 on PE-5 is as follows:

```
*A:PE-5# show router 1 pim group detail
```



```

=====
PIM Source Group ipv4
=====
Group Address      : 239.255.0.1
Source Address     : 172.16.5.2
RP Address         : 0
Advt Router        : 192.0.2.5
Flags              :
Mode               : sparse
MRIB Next Hop     : 172.16.5.2
MRIB Src Flags     : direct
Keepalive Timer    : Not Running
Up Time            : 0d 00:06:54
Resolved By       : rtable-u

Up JP State        : Joined
Up JP Rpt          : Not Joined StarG
Up JP Expiry       : 0d 00:00:00
Up JP Rpt Override : 0d 00:00:00

Register State     : No Info
Reg From Anycast RP: No

Rpf Neighbor       : 172.16.5.2
Incoming Intf      : int-PE-5-VPRN-1-S-5
Outgoing Intf List : mpls-if-73728

Curr Fwding Rate   : 78.560 kbps
Forwarded Packets  : 4140
Forwarded Octets   : 4065480
Spt threshold      : 0 kbps
Admin bandwidth    : 1 kbps
Discarded Packets  : 0
RPF Mismatches     : 0
ECMP opt threshold : 7

-----
Groups : 1
=====

```

## Conclusion

Inter-AS model C multicast within a VPRN can be achieved using non-segmented mLDP trees. This chapter provides the configuration for inter-AS model C MVPN. The example also shows the associated commands, debug, and outputs, which can be used for verifying and troubleshooting.

## NG-MVPN Sender-Only, Receiver-Only

This chapter provides information about next generation multicast virtual private network (NG-MVPN) sender-only and receiver-only configurations.

Topics in this chapter include:

- [Applicability](#)
- [Overview](#)
- [Configuration](#)
- [Conclusion](#)

### Applicability

The sender-only/receiver-only feature as described in this chapter is supported in SR OS Release 11.0.R1, and later. The CLI in this edition is based on SR OS Release 23.10.R1.

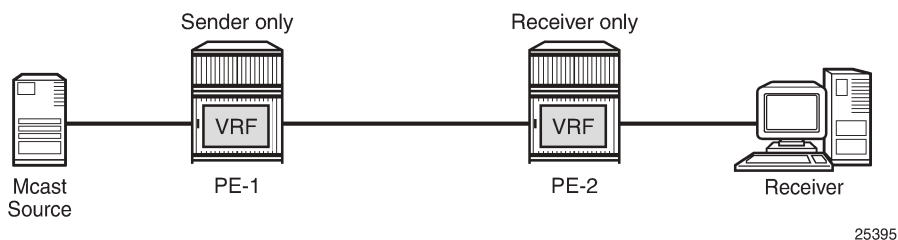
Knowledge of the Nokia multicast and Layer 3 VPNs concepts are assumed throughout this document.

### Overview

This example covers a basic technology overview, the network topology, and configuration examples which are used for the Multicast VPN (MVPN) sender-only, receiver-only feature.

By default, if multiple PE nodes form a peering relationship within a common MVPN instance, then each PE node originates a multicast tree locally toward the remaining PE nodes that are members of this MVPN instance. This behavior creates a full mesh of Inclusive-Provider Multicast Service Interfaces (I-PMSIs) across all PE nodes in the MVPN.

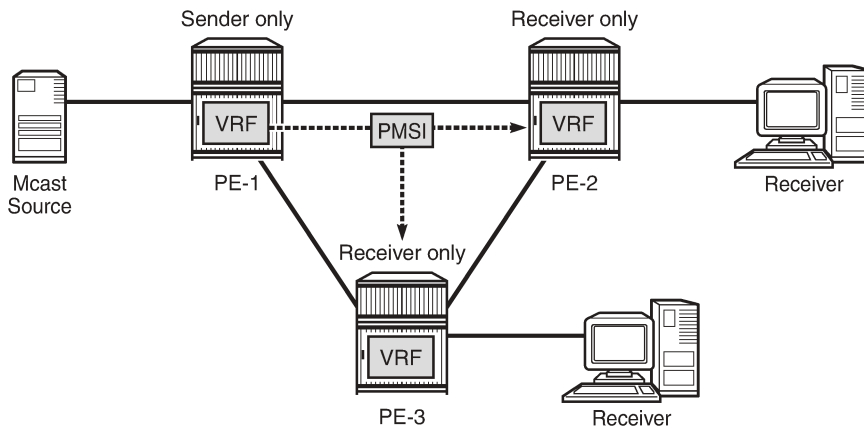
*Figure 52: Default PMSI Hierarchy*



It is often the case that an MVPN has many sites with multicast receivers, but only a few sites that host either both receivers and sources, or sources only.

The MVPN sender-only/receiver-only feature optimizes control and data plane resources by preventing unnecessary I-PMSI meshing when a PE hosts multicast sources only, or multicast receivers only, for an MVPN. An example of such an optimization is presented in [Figure 53: Optimized PMSI Structure](#).

Figure 53: Optimized PMSI Structure



The general rules to follow are:

- For PE nodes that host only multicast sources for a specific MVPN, operators can now block these PEs, through configuration, from joining I-PMSIs from the other PEs in this MVPN.
- For PE nodes that host only multicast receivers for a specific MVPN, operators can now block these PEs, through configuration, to set-up a local I-PMSI to the other PEs in this MVPN.

MVPN sender-only/receiver-only is supported with next generation MVPN for both IPv4 and IPv6 customer multicast using:

- IPv4 RSVP-TE provider tunnels
- IPv4 LDP provider tunnels

Extra attention should be given to the Bootstrap Router/Rendezvous Point (BSR/RP) placement when sender-only/receiver-only is enabled:

- The RP should be at the sender-receiver or sender-only site so that (\*,G) traffic can be sent over the tunnel
- The BSR should be deployed at the sender-receiver site.
- The BSR can be at a sender-only site if the RPs are at the same site.



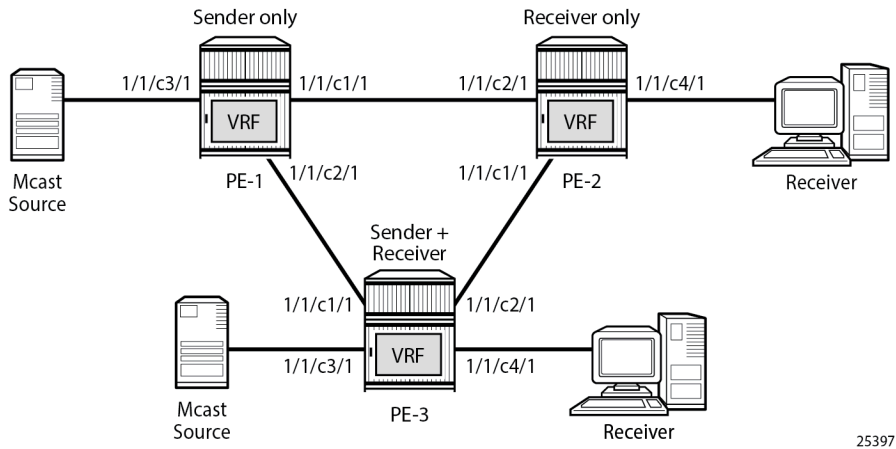
**Note:**

(\*,G) refers to an individual multicast stream indicating any source (\*) and the multicast group (G) used by the stream.

## Configuration

The example topology is shown in [Figure 54: Example Topology](#).

Figure 54: Example Topology



To configure the sender-only/receiver-only feature, the following configuration command is used:

```
# on PE-1
configure
service
  vprn 1 name "VPRN 1" customer 1 create
  mvpn
    mdt-type ?
  - mdt-type {sender-only|receiver-only|sender-receiver}
  - no mdt-type
```

**sender-receiver** is the default option and is visible using the **info detail** command.

This command restricts the MVPN instance to a specific role and provides an option to configure either a sender-only or receiver-only mode per PE node per service.

Parameters:

**sender-only** — MVPN has only senders connected to PE node.

**receiver-only** — MVPN has only receivers connected to PE node.

**sender-receiver** — MVPN has both sender and receivers connected to PE node.

Considerations:

- Two general approaches for building MVPNs are covered in detail in this example:
  - Point-to-multipoint (P2MP) RSVP MVPNs
  - Multicast LDP (mLDP) MVPNs
- IPv4 and IPv6 multicast streaming are used for every MVPN at the same time.
- Basic principles of an MVPN including I-PMSI, S-PMSI, mLDP and P2MP RSVP are covered in the [NG-MVPN Configuration with PIM](#) and [NG-MVPN Configuration with MPLS](#) chapters of this guide.

PIM SSM is used for IPv4/IPv6 Customer (C)-multicast groups.

## Initial Configuration

Step 1: The PE routers already have the following configuration:

- Interfaces (IPv4/IPv6)
- IGP (IS-IS or OSPF/OSPFv3)
- LDP (IPv4 only suffices)
- MPLS/RSVP
- BGP

Step 2: The MPLS/RSVP configuration on PE-1 is as follows. An P2MP LSP template is created with an empty path, without explicit hops.

```
# on PE-1
configure
router
  mpls
    interface "system"
      no shutdown
    exit
    interface "int-PE-1-PE-2"
      no shutdown
    exit
    interface "int-PE-1-PE-3"
      no shutdown
    exit
    path "EMPTY"
      no shutdown
    exit
    lsp-template "MVPN-P2MP-LSP" p2mp
      default-path "EMPTY"
      path-computation-method local-cspf
      no shutdown
    exit
  no shutdown
exit
  rsvp no shutdown
exit all
```

Step 3: The BGP configuration on PE-1 is as follows. No route reflector is used.

```
# on PE-1
configure
router
  bgp
    enable-peer-tracking
    rapid-withdrawal
    rapid-update mvpn-ipv4 mvpn-ipv6
    group "INTERNAL"
      family vpn-ipv4 vpn-ipv6 mvpn-ipv4 mvpn-ipv6
      type internal
      neighbor 192.0.2.2
      exit
      neighbor 192.0.2.3
      exit
    exit
  no shutdown
exit all
```

## RSVP-Based MVPN Configuration

Step 1: Configure a basic MVPN using P2MP RSVP as a transport protocol for C-multicast groups.

For this setup, PE-2 and PE-3 are configured to receive the following multicast groups:

- IPv4 group 232.0.0.1, source 172.16.1.2
- IPv6 group FF3E::8000:1, source 2001:DB8:1::2

Step 2: Configure the MDT type for the MVPN.

Based on the example topology, PE-1 is configured as **sender-only** for the MVPN.

```
# on PE-1
configure
  service
    vprn 1 name "VPRN 1" customer 1 create
      description "RSVP-based MVPN"
      ecmp 2
      autonomous-system 64500
      ignore-nh-metric
      interface "int-PE-1-S-1" create
        description "to multicast source S-1"
        address 172.16.1.2/30
        ipv6
          address 2001:db8:1::2/126
        exit
        sap 1/1/c3/1 create
        exit
      exit
    bgp-ipvpn
      mpls
        auto-bind-tunnel
        resolution-filter
        ldp
        rsvp
        exit
        resolution filter
        exit
        route-distinguisher 64500:101
        vrf-target target:64500:1
        no shutdown
      exit
    exit
  pim
    no ipv6-multicast-disable
    apply-to all
    no shutdown
  exit
  mvpn
    auto-discovery default
    c-mcast-signaling bgp
    mdt-type sender-only
    provider-tunnel
      inclusive
      rsvp
        lsp-template "MVPN-P2MP-LSP"
        no shutdown
      exit
    exit
  exit
  vrf-target unicast
```

```

    exit
  exit
  no shutdown
exit all

```

Based on the example topology, PE-2 is configured as **receiver-only** for the MVPN. PE-2 has static joins for the IPv4 and IPv6 multicast groups:

- group 232.0.0.1, source 172.16.1.2
- group FF3E::8000:1, source 2001:DB8:1::2

```

# on PE-2
configure
service
  vprn 1 name "VPRN 1" customer 1 create
  description "RSVP-based MVPN"
  ecmp 2
  autonomous-system 64500
  ignore-nh-metric
  interface "int-PE-2-H-2" create
    description "to receiver Host-2"
    address 172.16.2.2/30
    ipv6
      address 2001:db8:2::2/126
    exit
    sap 1/1/c4/1 create
    exit
  exit
  bgp-ipvpn
  mpls
    auto-bind-tunnel
    resolution-filter
    ldp
    rsvp
    exit
    resolution filter
  exit
  route-distinguisher 64500:102
  vrf-target target:64500:1
  no shutdown
  exit
exit
igmp
  interface "int-PE-2-H-2"
  static
    group 232.0.0.1
    source 172.16.1.2
  exit
  exit
  no shutdown
exit
no shutdown
exit
mld
  interface "int-PE-2-H-2"
  static
    group ff3e::8000:1
    source 2001:db8:1::2
  exit
  exit
  no shutdown
exit
no shutdown

```

```

exit
pim
  no ipv6-multicast-disable
  apply-to all
  no shutdown
exit
mvpn
  auto-discovery default
  c-mcast-signaling bgp
  mdt-type receiver-only
  provider-tunnel
    inclusive
    rsvp
      lsp-template "MVPN-P2MP-LSP"
      no shutdown
    exit
  exit
exit
vrf-target unicast
exit
exit
no shutdown
exit all

```

Based on the example topology, PE-3 is configured as **sender-receiver** (default) for the MVPN. PE-3 has also static joins for the IPv4 and IPv6 multicast groups:

- group 232.0.0.1, source 172.16.1.2
- group FF3E::8000:1, source 2001:DB8:1::2

The interface to the local source for PE-3 is not configured in this example. PE-3 acts as a receiver, not as a sender. Nonetheless, it is configured as sender-receiver and that has its consequences for the I-PMSIs that are established.

```

# on PE-3
configure
  service
    vprn 1 name "VPRN 1" customer 1 create
      description "RSVP-based MVPN"
      ecmp 2
      autonomous-system 64500
      ignore-nh-metric
      interface "int-PE-3-H-3" create
        description "to receiver Host-3"
        address 172.16.3.2/30
        ipv6
          address 2001:db8:3::2/126
        exit
        sap 1/1/c4/1 create
        exit
      exit
    bgp-ipvpn
      mpls
        auto-bind-tunnel
        resolution-filter
          ldp
          rsvp
        exit
        resolution filter
      exit
    route-distinguisher 64500:103
    vrf-target target:64500:1

```



```

        no shutdown
    exit
exit
igmp
    interface "int-PE-3-H-3"
        static
            group 232.0.0.1
            source 172.16.1.2
        exit
    exit
    no shutdown
exit
no shutdown
exit
mld
    interface "int-PE-3-H-3"
        static
            group ff3e::8000:1
            source 2001:db8:1::2
        exit
    exit
    no shutdown
exit
no shutdown
exit
pim
    no ipv6-multicast-disable
    apply-to all
    no shutdown
exit
mvpn
    auto-discovery default
    c-mcast-signaling bgp
    provider-tunnel
        inclusive
        rsvp
            lsp-template "MVPN-P2MP-LSP"
            no shutdown
        exit
    exit
    vrf-target unicast
exit
no shutdown
exit all

```

The PIM instance must be **shutdown** before the mdt-type is modified; this leads to a multicast service disruption. Trying to change the mdt-type with PIM instance active results in the following message being displayed.

```

# on PE-1
configure service vprn 1 mvpn mdt-type receiver-only
MINOR: PIM #1100 PIM instance must be shutdown before changing this configuration

```

## RSVP-Based MVPN Verification and Debugging

### MDT-Type Verification

The status of the MVPN can be checked using the **show router <service-number> mvpn** command:

The output for PE-1, PE-2 and PE-3 is as follows:

```
*A:PE-1# show router 1 mvpn

=====
MVPN 1 configuration data
=====
signaling          : Bgp                auto-discovery      : Default
UMH Selection      : Highest-Ip           SA withdrawn        : Disabled
intersite-shared   : Enabled                Persist SA          : Disabled
vrf-import         : N/A
vrf-export         : N/A
vrf-target         : unicast
C-Mcast Import RT : target:192.0.2.1:2

ipmsi              : rsvp MVPN-P2MP-LSP
i-pmsi P2MP AdmSt  : Up
i-pmsi Tunnel Name : MVPN-P2MP-LSP-1-73728
enable-bfd-root    : false                enable-bfd-leaf     : false
Mdt-type           : sender-only
ipmsi UMH RM       : Disabled

BSR signalling     : none
Wildcard s-pmsi   : Disabled
Multistream-SPMSI : Disabled
s-pmsi             : none
data-delay-interval : 3 seconds
enable-asm-mdt    : N/A
spmsi UMH RM      : Disabled

=====
```

```
*A:PE-2# show router 1 mvpn

=====
MVPN 1 configuration data
=====
signaling          : Bgp                auto-discovery      : Default
UMH Selection      : Highest-Ip           SA withdrawn        : Disabled
intersite-shared   : Enabled                Persist SA          : Disabled
vrf-import         : N/A
vrf-export         : N/A
vrf-target         : unicast
C-Mcast Import RT : target:192.0.2.2:2

ipmsi              : rsvp MVPN-P2MP-LSP
i-pmsi P2MP AdmSt  : Up
i-pmsi Tunnel Name : mpls-virt-if-1005857
enable-bfd-root    : false                enable-bfd-leaf     : false
Mdt-type           : receiver-only
ipmsi UMH RM       : Disabled

BSR signalling     : none
```

```

Wildcard s-pmsi      : Disabled
Multistream-SPMSI   : Disabled
s-pmsi              : none
data-delay-interval: 3 seconds
enable-asm-mdt      : N/A
spmsi UMH RM        : Disabled
=====

```

```
*A:PE-3# show router 1 mvpn
```

```
=====
MVPN 1 configuration data
=====
```

```

signaling           : Bgp                auto-discovery      : Default
UMH Selection       : Highest-Ip         SA withdrawn        : Disabled
intersite-shared    : Enabled            Persist SA          : Disabled
vrf-import          : N/A
vrf-export          : N/A
vrf-target          : unicast
C-Mcast Import RT  : target:192.0.2.3:2

```

```

ipmsi               : rsvp MVPN-P2MP-LSP
i-pmsi P2MP AdmSt   : Up
i-pmsi Tunnel Name  : MVPN-P2MP-LSP-1-73728
enable-bfd-root     : false              enable-bfd-leaf     : false
Mdt-type          : sender-receiver
ipmsi UMH RM        : Disabled

```

```

BSR signalling      : none
Wildcard s-pmsi     : Disabled
Multistream-SPMSI  : Disabled
s-pmsi             : none
data-delay-interval: 3 seconds
enable-asm-mdt     : N/A
spmsi UMH RM       : Disabled
=====

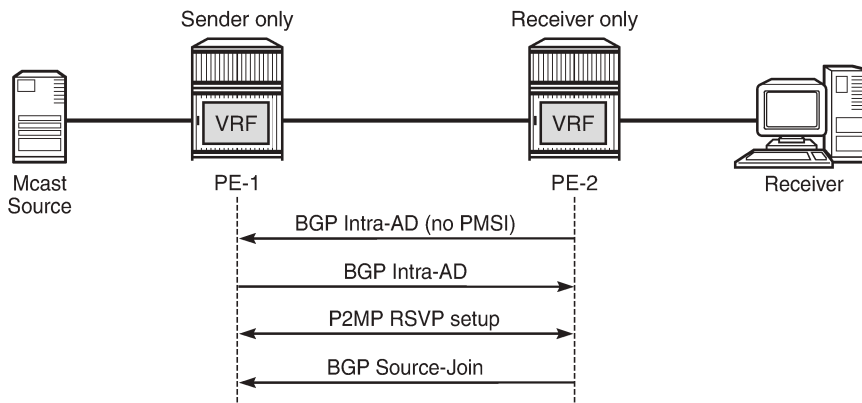
```

## BGP Verification and Debugging

When the MDT type is changed, the BGP signaling is slightly modified to achieve the signaling optimization.

The PE router does not include the PMSI part in the Intra-AD BGP messages when the MVPN is configured with **mdt-type receiver-only**. The message flow is presented in [Figure 55: RSVP-Based BGP Message Flow Between PE-1 and PE-2](#).

Figure 55: RSVP-Based BGP Message Flow Between PE-1 and PE-2



25398

The following BGP debug output is taken from PE-2 and demonstrates the message flow between PE-1 and PE-2 for the MVPN-IPv4 address family.

There is no PMSI part in the BGP Intra-AD message sent by PE-2 (message 6), but the PMSI part is present in the BGP Intra-AD message received from **sender-only** PE-1 (message 1).

```

1 2023/11/09 13:45:32.874 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Received BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 86
  Flag: 0x90 Type: 14 Len: 23 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.1
    Type: Intra-AD Len: 12 RD: 64500:101 Orig: 192.0.2.1
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:64500:1
  Flag: 0xc0 Type: 22 Len: 17 PMSI:
    Tunnel-type RSVP-TE P2MP LSP (1)
    Flags: (0x0)[Type: None BM: 0 U: 0 Leaf: not required]
    MPLS Label 0
    P2MP-ID 0x1, Tunnel-ID: 61441, Extended-Tunnel-ID 192.0.2.1
"

```

```

6 2023/11/09 13:45:45.463 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 66
  Flag: 0x90 Type: 14 Len: 23 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.2
    Type: Intra-AD Len: 12 RD: 64500:102 Orig: 192.0.2.2
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:

```

```
Flag: 0x80 Type: 4 Len: 4 MED: 0
Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:64500:1
"

19 2023/11/09 13:45:49.023 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Send BGP UPDATE:
    Withdrawn Length = 0
    Total Path Attr Length = 76
    Flag: 0x90 Type: 14 Len: 33 Multiprotocol Reachable NLRI:
        Address Family MVPN_IPV4
        NextHop len 4 NextHop 192.0.2.2
        Type: Source-Join Len:22 RD: 64500:101 SrcAS: 64500 Src: 172.16.1.2 Grp: 232.0.0.1
    Flag: 0x40 Type: 1 Len: 1 Origin: 0
    Flag: 0x40 Type: 2 Len: 0 AS Path:
    Flag: 0x80 Type: 4 Len: 4 MED: 0
    Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
    Flag: 0xc0 Type: 8 Len: 4 Community:
        no-export
    Flag: 0xc0 Type: 16 Len: 8 Extended Community:
        target:192.0.2.1:2
"
```

Similar behavior is observed for IPv6 multicast. The following BGP debug output is also taken from PE-2 and demonstrates the message flow between PE-1 and PE-2 for the MVPN-IPv6 address family.

There is no PMSI part in the Intra-AD message sent by PE-2 (message 8).

```
2 2023/11/09 13:45:32.874 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Received BGP UPDATE:
    Withdrawn Length = 0
    Total Path Attr Length = 86
    Flag: 0x90 Type: 14 Len: 23 Multiprotocol Reachable NLRI:
        Address Family MVPN_IPV6
        NextHop len 4 NextHop 192.0.2.1
        Type: Intra-AD Len: 12 RD: 64500:101 Orig: 192.0.2.1
    Flag: 0x40 Type: 1 Len: 1 Origin: 0
    Flag: 0x40 Type: 2 Len: 0 AS Path:
    Flag: 0x80 Type: 4 Len: 4 MED: 0
    Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
    Flag: 0xc0 Type: 8 Len: 4 Community:
        no-export
    Flag: 0xc0 Type: 16 Len: 8 Extended Community:
        target:64500:1
    Flag: 0xc0 Type: 22 Len: 17 PMSI:
        Tunnel-type RSVP-TE P2MP LSP (1)
        Flags: (0x0)[Type: None BM: 0 U: 0 Leaf: not required]
        MPLS Label 0
        P2MP-ID 0x1, Tunnel-ID: 61441, Extended-Tunnel-ID 192.0.2.1
"
```

```
8 2023/11/09 13:45:45.463 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Send BGP UPDATE:
    Withdrawn Length = 0
    Total Path Attr Length = 66
    Flag: 0x90 Type: 14 Len: 23 Multiprotocol Reachable NLRI:
```

```

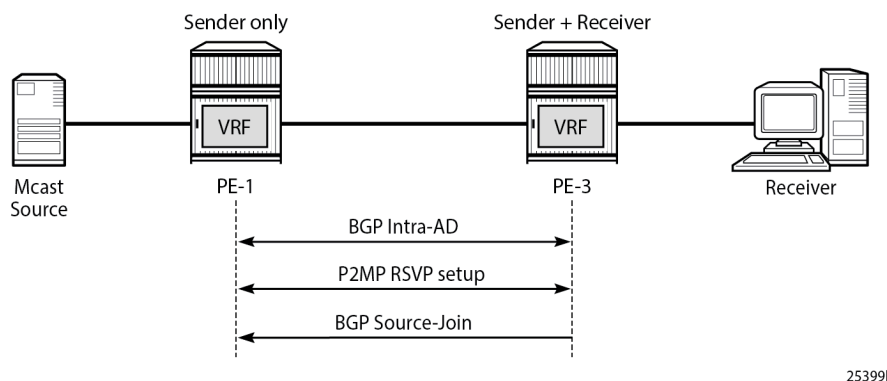
Address Family MVPN_IPV6
NextHop len 4 NextHop 192.0.2.2
Type: Intra-AD Len: 12 RD: 64500:102 Orig: 192.0.2.2
Flag: 0x40 Type: 1 Len: 1 Origin: 0
Flag: 0x40 Type: 2 Len: 0 AS Path:
Flag: 0x80 Type: 4 Len: 4 MED: 0
Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
Flag: 0xc0 Type: 8 Len: 4 Community:
no-export
Flag: 0xc0 Type: 16 Len: 8 Extended Community:
target:64500:1
"

20 2023/11/09 13:45:49.023 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Send BGP UPDATE:
Withdrawn Length = 0
Total Path Attr Length = 100
Flag: 0x90 Type: 14 Len: 57 Multiprotocol Reachable NLRI:
Address Family MVPN_IPV6
NextHop len 4 NextHop 192.0.2.2
Type: Source-Join Len:46 RD: 64500:101 SrcAS: 64500 Src: 2001:db8:1::2 Grp:
ff3e::8000:1
Flag: 0x40 Type: 1 Len: 1 Origin: 0
Flag: 0x40 Type: 2 Len: 0 AS Path:
Flag: 0x80 Type: 4 Len: 4 MED: 0
Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
Flag: 0xc0 Type: 8 Len: 4 Community:
no-export
Flag: 0xc0 Type: 16 Len: 8 Extended Community:
target:192.0.2.1:2
"

```

The PE router does not change its BGP behavior when the MVPN is configured with **mdt-type sender-only**. The message flow is presented in [Figure 56: RSVP-Based BGP Message Flow Between PE-1 and PE-3](#).

Figure 56: RSVP-Based BGP Message Flow Between PE-1 and PE-3



The BGP following debug output is taken from PE-3 and demonstrates the message flow between PE-1 and PE-3 for the MVPN-IPv4 address family.

The PMSI part is present in debug message 1, which is sent by PE-1 (**sender-only**).

```

1 2023/11/09 13:45:32.392 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1

```

```
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Received BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 86
  Flag: 0x90 Type: 14 Len: 23 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.1
    Type: Intra-AD Len: 12 RD: 64500:101 Orig: 192.0.2.1
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:64500:1
  Flag: 0xc0 Type: 22 Len: 17 PMSI:
    Tunnel-type RSVP-TE P2MP LSP (1)
    Flags: (0x0)[Type: None BM: 0 U: 0 Leaf: not required]
    MPLS Label 0
    P2MP-ID 0x1, Tunnel-ID: 61441, Extended-Tunnel-ID 192.0.2.1
"
```

```
13 2023/11/09 13:45:56.790 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 86
  Flag: 0x90 Type: 14 Len: 23 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.3
    Type: Intra-AD Len: 12 RD: 64500:103 Orig: 192.0.2.3
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:64500:1
  Flag: 0xc0 Type: 22 Len: 17 PMSI:
    Tunnel-type RSVP-TE P2MP LSP (1)
    Flags: (0x0)[Type: None BM: 0 U: 0 Leaf: not required]
    MPLS Label 0
    P2MP-ID 0x1, Tunnel-ID: 61441, Extended-Tunnel-ID 192.0.2.3
"
```

```
27 2023/11/09 13:46:01.060 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 76
  Flag: 0x90 Type: 14 Len: 33 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.3
    Type: Source-Join Len:22 RD: 64500:101 SrcAS: 64500 Src: 172.16.1.2 Grp: 232.0.0.1
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
"
```

```
Flag: 0xc0 Type: 16 Len: 8 Extended Community:
target:192.0.2.1:2
"
```

Similar behavior is observed for IPv6 multicast.

The following BGP debug output is taken from PE-3 and demonstrates the message flow between PE-1 and PE-3 for the MVPN-IPv6 address family.

The PMSI part is present in debug message 2, which is sent by PE-1 (**sender-only**).

```
2 2023/11/09 13:45:32.392 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Received BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 86
  Flag: 0x90 Type: 14 Len: 23 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV6
    NextHop len 4 NextHop 192.0.2.1
    Type: Intra-AD Len: 12 RD: 64500:101 Orig: 192.0.2.1
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:64500:1
  Flag: 0xc0 Type: 22 Len: 17 PMSI:
    Tunnel-type RSVP-TE P2MP LSP (1)
    Flags: (0x0)[Type: None BM: 0 U: 0 Leaf: not required]
    MPLS Label 0
    P2MP-ID 0x1, Tunnel-ID: 61441, Extended-Tunnel-ID 192.0.2.1
"
```

```
14 2023/11/09 13:45:56.790 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 86
  Flag: 0x90 Type: 14 Len: 23 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV6
    NextHop len 4 NextHop 192.0.2.3
    Type: Intra-AD Len: 12 RD: 64500:103 Orig: 192.0.2.3
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:64500:1
  Flag: 0xc0 Type: 22 Len: 17 PMSI:
    Tunnel-type RSVP-TE P2MP LSP (1)
    Flags: (0x0)[Type: None BM: 0 U: 0 Leaf: not required]
    MPLS Label 0
    P2MP-ID 0x1, Tunnel-ID: 61441, Extended-Tunnel-ID 192.0.2.3
"
```

```
28 2023/11/09 13:46:01.060 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Send BGP UPDATE:
```



```

Withdrawn Length = 0
Total Path Attr Length = 100
Flag: 0x90 Type: 14 Len: 57 Multiprotocol Reachable NLRI:
  Address Family MVPN_IPV6
  NextHop len 4 NextHop 192.0.2.3
  Type: Source-Join Len:46 RD: 64500:101 SrcAS: 64500 Src: 2001:db8:1::2 Grp:
ff3e::8000:1
Flag: 0x40 Type: 1 Len: 1 Origin: 0
Flag: 0x40 Type: 2 Len: 0 AS Path:
Flag: 0x80 Type: 4 Len: 4 MED: 0
Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
Flag: 0xc0 Type: 8 Len: 4 Community:
  no-export
Flag: 0xc0 Type: 16 Len: 8 Extended Community:
  target:192.0.2.1:2
"

```

The BGP routing table of each router is populated accordingly.

PE-1 (**sender-only**) has two Intra-Ad and two Source-Join messages from PE-2 and PE-3.

```

*A:PE-1# show router bgp routes mvpn-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 MVPN-IPv4 Routes
=====
Flag RouteType      OriginatorIP      LocalPref  MED
      RD            SourceAS          Path-Id     IGP Cost
      Nexthop      SourceIP          Label
      As-Path      GroupIP
-----
u*>i Source-Join      -                100        0
      64500:101      64500            None        -
      192.0.2.2      172.16.1.2
      No As-Path    232.0.0.1
*>i Source-Join      -                100        0
      64500:101      64500            None        -
      192.0.2.3      172.16.1.2
      No As-Path    232.0.0.1
u*>i Intra-Ad        192.0.2.2        100        0
      64500:102      -                None        -
      192.0.2.2      -
      No As-Path    -
u*>i Intra-Ad        192.0.2.3        100        0
      64500:103      -                None        -
      192.0.2.3      -
      No As-Path    -
-----
Routes : 4
=====

```

PE-2 (**receiver-only**) has two Intra-Ad messages from PE-1 and PE-3.

```

*A:PE-2# show router bgp routes mvpn-ipv4
=====
BGP Router ID:192.0.2.2      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 MVPN-IPv4 Routes

```

=====
Flag  RouteType      OriginatorIP      LocalPref  MED
      RD            SourceAS          Path-Id     IGP Cost
      Nexthop       SourceIP          Label
      As-Path      GroupIP
-----
u*>i  Intra-Ad          192.0.2.1        100        0
      64500:101      -                None        -
      192.0.2.1     -
      No As-Path    -
u*>i  Intra-Ad          192.0.2.3        100        0
      64500:103      -                None        -
      192.0.2.3     -
      No As-Path    -
=====

```

Routes : 2

PE-3 (**sender-receiver**) has two Intra-Ad messages: one from PE-1 and one from PE-2.

```
*A:PE-3# show router bgp routes mvpn-ipv4
```

```

=====
BGP Router ID:192.0.2.3      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 MVPN-IPv4 Routes

```

=====
Flag  RouteType      OriginatorIP      LocalPref  MED
      RD            SourceAS          Path-Id     IGP Cost
      Nexthop       SourceIP          Label
      As-Path      GroupIP
-----
u*>i  Intra-Ad          192.0.2.1        100        0
      64500:101      -                None        -
      192.0.2.1     -
      No As-Path    -
u*>i  Intra-Ad          192.0.2.2        100        0
      64500:102      -                None        -
      192.0.2.2     -
      No As-Path    -
=====

```

Routes : 2

## RSVP Verification and Debugging

When BGP intra-AD messages are exchanged, every PE starts to build multicast tunnels based on the following criteria:

- PE nodes which are configured as **sender-only** for an MVPN do not join P2MP LSPs from other PEs in this MVPN.
- PE nodes which are configured as **receiver-only** for an MVPN do not originate P2MP LSPs to other PEs in this MVPN.

The RSVP session can be checked with the **show router rsvp session** command:

PE-1 (192.0.2.1) has two originating LSPs: one toward PE-2 (192.0.2.2) and one toward PE-3 (192.0.2.3). PE-1 also has one incoming LSP from PE-3 (**mdt-type sender-receiver**).

```
*A:PE-1# show router rsvp session
=====
RSVP Sessions
=====
RSVP Session Name
  From           To           Tunnel ID   LSP ID      State
-----
MVPN-P2MP-LSP-1-73728::EMPTY
192.0.2.1       192.0.2.2   61441      26112      Up
MVPN-P2MP-LSP-1-73728::EMPTY
192.0.2.1       192.0.2.3   61441      26112      Up
MVPN-P2MP-LSP-1-73728::EMPTY
192.0.2.3       192.0.2.1   61441      58880      Up
-----
Sessions : 3
=====
```

PE-2 (192.0.2.2) has two incoming LSPs from PE-1 (192.0.2.1) and PE-3 (192.0.2.3) and no originating LSPs because PE-2 has **mdt-type receiver-only**.

```
*A:PE-2# show router rsvp session
=====
RSVP Sessions
=====
RSVP Session Name
  From           To           Tunnel ID   LSP ID      State
-----
MVPN-P2MP-LSP-1-73728::EMPTY
192.0.2.1       192.0.2.2   61441      26112      Up
MVPN-P2MP-LSP-1-73728::EMPTY
192.0.2.3       192.0.2.2   61441      58880      Up
-----
Sessions : 2
=====
```

PE-3 (192.0.2.3) has two originating LSPs: one toward PE-2 (192.0.2.2) and one toward PE-1 (192.0.2.1). PE-3 also has one incoming LSP from PE-1 (**mdt-type sender-only**).

Theoretically there is no need for the LSP from PE-3 toward PE-1, because PE-1 is a sender-only; this minor limitation should be taken into account during planning phase.

```
*A:PE-3# show router rsvp session

=====
RSVP Sessions
=====
RSVP Session Name
  From           To           Tunnel ID   LSP ID     State
-----
MVPN-P2MP-LSP-1-73728::EMPTY
192.0.2.1       192.0.2.3       61441      26112     Up
MVPN-P2MP-LSP-1-73728::EMPTY
192.0.2.3       192.0.2.1       61441      58880     Up
MVPN-P2MP-LSP-1-73728::EMPTY
192.0.2.3       192.0.2.2       61441      58880     Up
-----
Sessions : 3
=====
```

Additional details about originating P2MP paths can be found using the following command:

**show router mpls p2mp-lsp <lsp name> p2mp-instance <service number> s2l**

The output for PE-1, PE-2 and PE-3 is as follows:

```
*A:PE-1# show router mpls p2mp-lsp "MVPN-P2MP-LSP-1-73728" p2mp-instance "1" s2l

=====
MPLS LSP MVPN-P2MP-LSP-1-73728 S2L
=====
-----
LSP Name      : MVPN-P2MP-LSP-1-73728
P2MP ID       : 1
Adm State     : Up           Oper State    : Up
P2MPInstance  : 1
Inst-type     : Primary
Adm State     : Up           Oper State    : Up
-----
S2L Name      To           Next Hop     Adm Opr
-----
EMPTY         192.0.2.2   192.168.12.2 Up  Up
EMPTY         192.0.2.3   192.168.13.2 Up  Up
=====
```

```
*A:PE-2# show router mpls p2mp-lsp

=====
MPLS P2MP LSPs (Originating)
=====
-----
LSP Name      Tun   Fastfail  Adm Opr
                Id    Config
-----
No Matching Entries Found
```

```

=====
*A:PE-3# show router mpls p2mp-lsp "MVPN-P2MP-LSP-1-73728" p2mp-instance "1" s2l
=====
MPLS LSP MVPN-P2MP-LSP-1-73728 S2L
=====
-----
LSP Name      : MVPN-P2MP-LSP-1-73728
P2MP ID       : 1
Adm State     : Up                Oper State     : Up
P2MPInstance  : 1
-----
Adm State     : Up                Inst-type      : Primary
Oper State    : Up
-----
S2L Name      To                Next Hop      Adm  Opr
-----
EMPTY
                192.0.2.1          192.168.13.1  Up   Up
EMPTY
                192.0.2.2          192.168.23.1  Up   Up
=====

```

## Multicast Stream Verification

The status of the multicast groups/streams can be verified using the **show router <sid> pim group detail [ipv6]** command:

There is an IPv4 sender connected to PE-1. The physical interface where the sender is connected is used as the incoming interface. An I-PMSI is used as the outgoing interface.

```

*A:PE-1# show router 1 pim group detail
=====
PIM Source Group ipv4
=====
Group Address      : 232.0.0.1
Source Address     : 172.16.1.2
RP Address         : 0
Advt Router       : 192.0.2.1
Flags              :
Mode              : sparse
MRIB Next Hop     : 172.16.1.2
MRIB Src Flags    : direct
Keepalive Timer   : Not Running
Up Time           : 0d 00:03:56
Resolved By       : rtable-u
Type              : (S,G)
Up JP State       : Joined
Up JP Rpt         : Not Joined StarG
Up JP Expiry      : 0d 00:00:00
Up JP Rpt Override : 0d 00:00:00
Register State    : No Info
Reg From Anycast RP: No
Rpf Neighbor      : 172.16.1.2
Incoming Intf     : int-PE-1-S-1
Outgoing Intf List : mpls-if-73728
Curr Fwding Rate  : 4811.800 kbps
Forwarded Packets : 31737
Forwarded Octets  : 31165734
Discarded Packets : 0
RPF Mismatches    : 0

```

```
Spt threshold      : 0 kbps          ECMP opt threshold : 7
Admin bandwidth   : 1 kbps
-----
Groups : 1
=====
```

There is an IPv4 receiver connected to PE-2. An I-PMSI is used as the incoming interface and the physical interface where the receiver is connected is used as the outgoing interface.

```
*A:PE-2# show router 1 pim group detail
=====
PIM Source Group ipv4
=====
Group Address      : 232.0.0.1
Source Address     : 172.16.1.2
RP Address         : 0
Advt Router        : 192.0.2.1
Flags              :                               Type           : (S,G)
Mode               : sparse
MRIB Next Hop     : 192.0.2.1
MRIB Src Flags    : remote
Keepalive Timer   : Not Running
Up Time           : 0d 00:04:05          Resolved By         : rtable-u

Up JP State       : Joined                Up JP Expiry        : 0d 00:00:58
Up JP Rpt         : Not Joined StarG      Up JP Rpt Override  : 0d 00:00:00

Register State    : No Info
Reg From Anycast RP: No

Rpf Neighbor      : 192.0.2.1
Incoming Intf   : mpls-if-73728
Outgoing Intf List : int-PE-2-H-2

Curr Fwding Rate  : 4811.800 kbps
Forwarded Packets : 35774                Discarded Packets   : 0
Forwarded Octets  : 35130068           RPF Mismatches      : 0
Spt threshold     : 0 kbps              ECMP opt threshold  : 7
Admin bandwidth   : 1 kbps
-----
Groups : 1
=====
```

There is an IPv4 receiver connected to PE-3. An I-PMSI is used as the incoming interface and the physical interface where receiver is connected is used as the outgoing interface.

```
*A:PE-3# show router 1 pim group detail
=====
PIM Source Group ipv4
=====
Group Address      : 232.0.0.1
Source Address     : 172.16.1.2
RP Address         : 0
Advt Router        : 192.0.2.1
Flags              :                               Type           : (S,G)
Mode               : sparse
MRIB Next Hop     : 192.0.2.1
MRIB Src Flags    : remote
Keepalive Timer   : Not Running
Up Time           : 0d 00:03:54          Resolved By         : rtable-u
```

```

Up JP State      : Joined          Up JP Expiry      : 0d 00:00:08
Up JP Rpt       : Not Joined StarG Up JP Rpt Override : 0d 00:00:00

Register State   : No Info
Reg From Anycast RP: No

Rpf Neighbor    : 192.0.2.1
Incoming Intf  : mpls-if-73729
Outgoing Intf List : int-PE-3-H-3

Curr Fwding Rate : 4815.728 kbps
Forwarded Packets : 36887
Forwarded Octets  : 36223034
Spt threshold    : 0 kbps
Admin bandwidth  : 1 kbps
Discarded Packets : 0
RPF Mismatches   : 0
ECMP opt threshold : 7
-----
Groups : 1
=====

```

Similar behavior is observed for IPv6 multicast.

An IPv6 sender is connected to PE-1. The physical interface where the sender is connected is used as the incoming interface. An I-PMSI is used as the outgoing interface.

```

*A:PE-1# show router 1 pim group detail ipv6
=====
PIM Source Group ipv6
=====
Group Address      : ff3e::8000:1
Source Address     : 2001:db8:1::2
RP Address         : 0
Advt Router       : 192.0.2.1
Flags              :
Type               : (S,G)
Mode               : sparse
MRIB Next Hop     : 2001:db8:1::2
MRIB Src Flags    : direct
Keepalive Timer   : Not Running
Up Time           : 0d 00:03:56
Resolved By       : rtable6-u

Up JP State      : Joined          Up JP Expiry      : 0d 00:00:00
Up JP Rpt       : Not Joined StarG Up JP Rpt Override : 0d 00:00:00

Register State   : No Info
Reg From Anycast RP: No

Rpf Neighbor    : 2001:db8:1::2
Incoming Intf  : int-PE-1-S-1
Outgoing Intf List : mpls-if-73728

Curr Fwding Rate : 4811.800 kbps
Forwarded Packets : 30721
Forwarded Octets  : 30168022
Spt threshold    : 0 kbps
Admin bandwidth  : 1 kbps
Discarded Packets : 0
RPF Mismatches   : 0
ECMP opt threshold : 7
-----
Groups : 1
=====

```

An IPv6 receiver is connected to PE-2. An I-PMSI is used as the incoming interface and the physical interface where the receiver is connected is used as the outgoing interface.

```
*A:PE-2# show router 1 pim group detail ipv6
=====
PIM Source Group ipv6
=====
Group Address      : ff3e::8000:1
Source Address     : 2001:db8:1::2
RP Address         : 0
Advt Router        : 192.0.2.1
Flags              :                               Type           : (S,G)
Mode               : sparse
MRIB Next Hop     : 192.0.2.1
MRIB Src Flags    : remote
Keepalive Timer   : Not Running
Up Time           : 0d 00:04:04      Resolved By         : rtable6-u

Up JP State       : Joined           Up JP Expiry        : 0d 00:00:58
Up JP Rpt        : Not Joined StarG  Up JP Rpt Override : 0d 00:00:00

Register State    : No Info
Reg From Anycast RP: No

Rpf Neighbor      : 192.0.2.1
Incoming Intf   : mpls-if-73728
Outgoing Intf List : int-PE-2-H-2

Curr Fwding Rate  : 4811.800 kbps
Forwarded Packets : 34759             Discarded Packets  : 0
Forwarded Octets  : 34133338        RPF Mismatches     : 0
Spt threshold     : 0 kbps           ECMP opt threshold : 7
Admin bandwidth   : 1 kbps
-----
Groups : 1
=====
```

An IPv6 receiver is connected to PE-3. An I-PMSI is used as the incoming interface and the physical interface where the receiver is connected is used as the outgoing interface.

```
*A:PE-3# show router 1 pim group detail ipv6
=====
PIM Source Group ipv6
=====
Group Address      : ff3e::8000:1
Source Address     : 2001:db8:1::2
RP Address         : 0
Advt Router        : 192.0.2.1
Flags              :                               Type           : (S,G)
Mode               : sparse
MRIB Next Hop     : 192.0.2.1
MRIB Src Flags    : remote
Keepalive Timer   : Not Running
Up Time           : 0d 00:03:53      Resolved By         : rtable6-u

Up JP State       : Joined           Up JP Expiry        : 0d 00:00:08
Up JP Rpt        : Not Joined StarG  Up JP Rpt Override : 0d 00:00:00

Register State    : No Info
Reg From Anycast RP: No
```



```
Rpf Neighbor      : 192.0.2.1
Incoming Intf    : mpls-if-73729
Outgoing Intf List : int-PE-3-H-3

Curr Fwding Rate  : 4811.800 kbps
Forwarded Packets : 35872
Forwarded Octets  : 35226304
Spt threshold    : 0 kbps
Admin bandwidth  : 1 kbps
Discarded Packets : 0
RPF Mismatches   : 0
ECMP opt threshold : 7
-----
Groups : 1
=====
```

## mLDP-Based MVPN Configuration

Step 1: Reconfigure VPRN 1 to make it mLDP-based. The resolution-filter should only be LDP (no RSVP anymore) for auto-bind-tunnel. The MVPN context also changes: the inclusive provider-tunnel is mLDP-based. The MDT-type remains the same: PE-1 is sender-only, PE-2 is receiver-only and PE-3 is sender-receiver (default).

PE-2 and PE-3 have static joins for the IPv4/IPv6 multicast groups:

- group 232.0.0.1, source 172.16.1.2
- group FF3E::8000:1, source 2001:DB8:1::2

Step 2: The VPRN 1 configuration on PE-1 is as follows:

```
# on PE-1
configure
  service
    vprn 1 name "VPRN 1" customer 1 create
      description "mLDP-based MVPN"
      ecmp 2
      autonomous-system 64500
      ignore-nh-metric
      interface "int-PE-1-S-1" create
        description "to multicast source S-1"
        address 172.16.1.2/30
        ipv6
          address 2001:db8:1::2/126
        exit
        sap 1/1/c3/1 create
        exit
      exit
    bgp-ipvpn
      mpls
        auto-bind-tunnel
        resolution-filter
          ldp
        exit
        resolution filter
      exit
      route-distinguisher 64500:101
      vrf-target target:64500:1
      no shutdown
    exit
  exit
  pim
    no ipv6-multicast-disable
```

```

        apply-to all
        no shutdown
    exit
    mvpn
        auto-discovery default
        c-mcast-signaling bgp
        mdt-type sender-only
        provider-tunnel
            inclusive
            mldp
                no shutdown
            exit
        exit
    exit
    vrf-target unicast
    exit
    exit
    no shutdown
    exit all

```

Based on the example topology, PE-2 is configured as receiver-only for the MVPN. PE-2 has also static joins for the IPv4 and IPv6 multicast groups:

- group 232.0.0.1, source 172.16.1.2
- group FF3E::8000:1, source 2001:DB8:1::2

```

# on PE-2
configure
  service
    vprn 1 name "VPRN 1" customer 1 create
      description "mLDP-based MVPN"
      ecmp 2
      autonomous-system 64500
      ignore-nh-metric
      interface "int-PE-2-H-2" create
        description "to receiver Host-2"
        address 172.16.2.2/30
        ipv6
          address 2001:db8:2::2/126
        exit
        sap 1/1/c4/1 create
        exit
      exit
    bgp-ipvpn
      mpls
        auto-bind-tunnel
        resolution-filter
          ldp
            exit
          resolution filter
        exit
        route-distinguisher 64500:102
        vrf-target target:64500:1
        no shutdown
      exit
    exit
  igmp
    interface "int-PE-2-H-2"
      static
        group 232.0.0.1
          source 172.16.1.2
        exit
      exit
    exit

```

```

        no shutdown
    exit
    no shutdown
exit
mld
    interface "int-PE-2-H-2"
        static
            group ff3e::8000:1
            source 2001:db8:1::2
        exit
    exit
    no shutdown
exit
no shutdown
exit
pim
    no ipv6-multicast-disable
    apply-to all
    no shutdown
exit
mvpn
    auto-discovery default
    c-mcast-signaling bgp
    mdt-type receiver-only
    provider-tunnel
        inclusive
        mldp
        no shutdown
    exit
    exit
    vrf-target unicast
exit
no shutdown
exit all

```

Based on the example topology, PE-3 is configured as **sender-receiver** (default) for the MVPN. PE-3 has also static joins for the IPv4 and IPv6 multicast groups:

- group 232.0.0.1, source 172.16.1.2
- group FF3E::8000:1, source 2001:DB8:1::2

```

# on PE-3
configure
    service
        vprn 1 name "VPRN 1" customer 1 create
            description "mLDP-based MVPN"
            ecmp 2
            autonomous-system 64500
            ignore-nh-metric
            interface "int-PE-3-H-3" create
                description "to receiver Host-3"
                address 172.16.3.2/30
                ipv6
                    address 2001:db8:3::2/126
                exit
                sap 1/1/c4/1 create
            exit
        exit
    bgp-ipvpn
        mpls
            auto-bind-tunnel

```

```
        resolution-filter
        ldp
        exit
        resolution filter
    exit
    route-distinguisher 64500:103
    vrf-target target:64500:1
    no shutdown
exit
exit
igmp
interface "int-PE-3-H-3"
    static
        group 232.0.0.1
        source 172.16.1.2
    exit
    exit
    no shutdown
exit
no shutdown
exit
mld
interface "int-PE-3-H-3"
    static
        group ff3e::8000:1
        source 2001:db8:1::2
    exit
    exit
    no shutdown
exit
no shutdown
exit
pim
no ipv6-multicast-disable
apply-to all
no shutdown
exit
mvpn
auto-discovery default
c-mcast-signaling bgp
provider-tunnel
    inclusive
    mldp
        no shutdown
    exit
    exit
    exit
    vrf-target unicast
    exit
exit
no shutdown
exit all
```

## mLDP-Based MVPN Verification and Debugging

### MDT-Type Verification

The status of the MVPN can be checked using the following command:

**show router** <service-number> **mvpn**

The output for PE-1, PE-2 and PE-3 is as follows:

```
*A:PE-1# show router 1 mvpn
```

```
=====
MVPN 1 configuration data
=====
```

```
signaling          : Bgp          auto-discovery     : Default
UMH Selection      : Highest-IP   SA withdrawn       : Disabled
intersite-shared   : Enabled      Persist SA         : Disabled
vrf-import         : N/A
vrf-export         : N/A
vrf-target         : unicast
C-Mcast Import RT : target:192.0.2.1:2
```

```
ipmsi              : ldp
i-pmsi P2MP AdmSt  : Up
i-pmsi Tunnel Name : mpls-if-73729
Mdt-type           : sender-only
ipmsi UMH RM       : Disabled
```

```
BSR signalling     : none
Wildcard s-pmsi   : Disabled
Multistream-SPMSI : Disabled
s-pmsi            : none
data-delay-interval: 3 seconds
enable-asm-mdt    : N/A
spmsi UMH RM      : Disabled
```

```
=====
```

```
*A:PE-2# show router 1 mvpn
```

```
=====
MVPN 1 configuration data
=====
```

```
signaling          : Bgp          auto-discovery     : Default
UMH Selection      : Highest-IP   SA withdrawn       : Disabled
intersite-shared   : Enabled      Persist SA         : Disabled
vrf-import         : N/A
vrf-export         : N/A
vrf-target         : unicast
C-Mcast Import RT : target:192.0.2.2:2
```

```
ipmsi              : ldp
i-pmsi P2MP AdmSt  : Up
i-pmsi Tunnel Name : mpls-virt-if-1005858
Mdt-type           : receiver-only
ipmsi UMH RM       : Disabled
```

```
BSR signalling     : none
Wildcard s-pmsi   : Disabled
Multistream-SPMSI : Disabled
s-pmsi            : none
data-delay-interval: 3 seconds
enable-asm-mdt    : N/A
spmsi UMH RM      : Disabled
```

```
=====
```

```
*A:PE-3# show router 1 mvpn
```

```

=====
MVPN 1 configuration data
=====
signaling          : Bgp          auto-discovery    : Default
UMH Selection      : Highest-Ip   SA withdrawn     : Disabled
intersite-shared   : Enabled      Persist SA       : Disabled
vrf-import         : N/A
vrf-export         : N/A
vrf-target         : unicast
C-Mcast Import RT : target:192.0.2.3:2

ipmsi              : ldp
i-pmsi P2MP AdmSt : Up
i-pmsi Tunnel Name: mpls-if-73730
Mdt-type           : sender-receiver
ipmsi UMH RM      : Disabled

BSR signalling     : none
Wildcard s-pmsi   : Disabled
Multistream-SPMSI: Disabled
s-pmsi            : none
data-delay-interval: 3 seconds
enable-asm-mdt    : N/A
spmsi UMH RM      : Disabled
=====

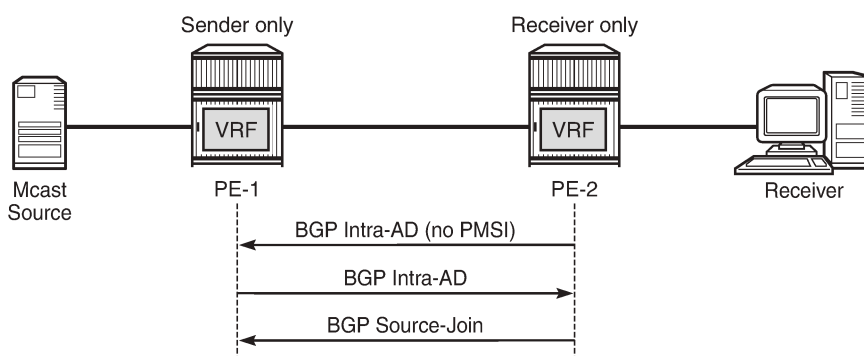
```

## BGP Verification and Debugging

When the MDT type is changed, the BGP signaling is slightly modified to achieve the signaling optimization. The PE router does not include the PMSI part in Intra-AD BGP messages when the MVPN is configured with **mdt-type receiver-only**.

The message flow is presented in [Figure 57: mLDP-Based BGP Message Flow Between PE-1 and PE-2](#).

Figure 57: mLDP-Based BGP Message Flow Between PE-1 and PE-2



25400

To demonstrate the BGP message flow sequence the following initialization steps are taken on PE-2:

1. Bring down the VPRN service, PIM protocol in a VPRN and IGMP/MLD protocol. As a result, the state of all signaling protocols is cleared.
2. Bring up the VPRN service. BGP exchanges unicast routing information.

3. Bring up the IPv4 PIM protocol. BGP exchanges IPv4 multicast routing information to build the PMSI infrastructure.
4. Bring up IGMP and add a static IGMP join where it is applicable. BGP exchanges IPv4 multicast routing information to propagate the multicast traffic to the receiver.
5. Bring up the IPv6 PIM protocol. BGP exchanges IPv6 multicast routing information to build the PMSI infrastructure.
6. Bring up MLD and add a static MLD join where it is applicable. BGP exchanges IPv6 multicast routing information to propagate the multicast traffic to the receiver.

The following BGP debug is taken from PE-2 and demonstrates the message flow between PE-2 and PE-1. VPN-IPv4 and VPN-IPv6 updates are not present in this output.

Step 1: Bring down the VPRN service and protocols to clear the state of all signaling protocols.

```
# on PE-2
configure
service
  vprn 1
  shutdown
  pim shutdown
  pim ipv6-multicast-disable
  igmp shutdown
  mld shutdown
exit all
```

Step 2: Enable the VPRN service on PE-2.

PE-2 immediately receives Intra-AD messages from PE-1 because the remote VPRN service is already enabled for IPv4 and IPv6 multicast propagation.

```
# on PE-2
configure service vprn 1 no shutdown
```

```
16 2023/11/09 13:55:33.021 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Received BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 91
  Flag: 0x90 Type: 14 Len: 23 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.1
    Type: Intra-AD Len: 12 RD: 64500:101 Orig: 192.0.2.1
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:64500:1
Flag: 0xc0 Type: 22 Len: 22 PMSI:
Tunnel-type LDP P2MP LSP (2)
Flags: (0x0)[Type: None BM: 0 U: 0 Leaf: not required]
MPLS Label 0
Root-Node 192.0.2.1, LSP-ID 0x2001
"
```

```
9 2023/11/09 13:55:33.020 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
```

```
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Received BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 91
  Flag: 0x90 Type: 14 Len: 23 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV6
    NextHop len 4 NextHop 192.0.2.1
    Type: Intra-AD Len: 12 RD: 64500:101 Orig: 192.0.2.1
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:64500:1
  Flag: 0xc0 Type: 22 Len: 22 PMSI:
    Tunnel-type LDP P2MP LSP (2)
    Flags: (0x0)[Type: None BM: 0 U: 0 Leaf: not required]
    MPLS Label 0
    Root-Node 192.0.2.1, LSP-ID 0x2001
"
```

Step 3: Enable only PIM IPv4 for the service on PE-2.

PE-2 immediately sends Intra-AD messages to PE-1. Note that no PMSI part is present in the debug message sent by receiver-only PE-2.

```
# on PE-2
configure service vprn 1 pim no shutdown
```

```
6 2023/11/09 13:55:32.205 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 66
  Flag: 0x90 Type: 14 Len: 23 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.2
    Type: Intra-AD Len: 12 RD: 64500:102 Orig: 192.0.2.2
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:64500:1
"
```

Step 4: Bring up IGMP and add a static IGMP join for the service on a PE-2.

PE-2 immediately sends a source-join message to PE-3 and receives a source-AD message from PE-1.

```
# on PE-2
configure
  service
    vprn 1
      igmp
        interface "int-PE-2-H-2"
          static
            group 232.0.0.1 source 172.16.1.2
```



```

        exit
        no shutdown
    exit
    no shutdown
exit
exit all

```

```

18 2023/11/09 13:55:38.214 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 76
  Flag: 0x90 Type: 14 Len: 33 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.2
    Type: Source-Join Len:22 RD: 64500:101 SrcAS: 64500 Src: 172.16.1.2 Grp: 232.0.0.1
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:192.0.2.1:2
"

```

Step 5: Enable PIM IPv6 for the service on PE-2.

PE-2 immediately sends Intra-AD messages to PE-3.

```

# on PE-2
configure service vprn 1 pim no ipv6-multicast-disable

```

```

20 2023/11/09 13:55:46.409 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 66
  Flag: 0x90 Type: 14 Len: 23 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV6
    NextHop len 4 NextHop 192.0.2.2
    Type: Intra-AD Len: 12 RD: 64500:102 Orig: 192.0.2.2
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:64500:1
"

```

Step 6: Bring up MLD and add a static MLD join for the service on a PE-2.

PE-2 immediately sends a source-join message to PE-3 and receives a source-AD message from PE-3.

```

# on PE-2
configure
  service
    vprn 1
      mld

```

```

interface "int-PE-2-H-2"
  static
  group FF3E::8000:1 source 2001:DB8:1::2
  exit
  no shutdown
exit
no shutdown
exit
exit all

```

```

22 2023/11/09 13:55:52.423 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 100
  Flag: 0x90 Type: 14 Len: 57 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV6
    NextHop len 4 NextHop 192.0.2.2
    Type: Source-Join Len:46 RD: 64500:101 SrcAS: 64500 Src: 2001:db8:1::2 Grp:
ff3e::8000:1
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:192.0.2.1:2
"

```

The same information can be gathered using the following show commands.

**show router bgp neighbor <peer> advertised-routes [mvpn-ipv4 | mvpn-ipv6]**

**show router bgp neighbor <peer> received-routes> [mvpn-ipv4 | mvpn-ipv6]**

PE-2 output for the advertised routes for the mvpn-ipv4 address family is as follows:

```

*A:PE-2# show router bgp neighbor 192.0.2.1 advertised-routes mvpn-ipv4
=====
BGP Router ID:192.0.2.2      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 MVPN-IPv4 Routes
=====
Flag  RouteType      OriginatorIP      LocalPref  MED
      RD             SourceAS          Path-Id    IGP Cost
      Nexthop        SourceIP          Label
      As-Path        GroupIP
-----
i     Source-Join     -                 100        0
      64500:101       64500            None       -
      192.0.2.2      172.16.1.2
      No As-Path     232.0.0.1
i     Intra-Ad       192.0.2.2        100        0
      64500:102       -                None       -
      192.0.2.2      -
      No As-Path     -

```

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

PE-2 output for the advertised routers for the mvpn-ipv6 address family is as follows:

```
*A:PE-2# show router bgp neighbor 192.0.2.1 advertised-routes mvpn-ipv6
=====
BGP Router ID:192.0.2.2      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 MVPN-IPv6 Routes
=====
Flag RouteType      OriginatorIP      LocalPref  MED
      RD            SourceAS          Path-Id     IGP Cost
      Nexthop       SourceIP          Label
      As-Path       GroupIP
-----
i     Source-Join    -                 100        0
      64500:101      64500            None        -
      192.0.2.2     2001:db8:1::2
      No As-Path    ff3e::8000:1
i     Intra-Ad      192.0.2.2        100        0
      64500:102     -                 None        -
      192.0.2.2     -
      No As-Path    -
-----
Routes : 2
=====
```

PE-2 output for the received routes for the mvpn-ipv4 address family is as follows:

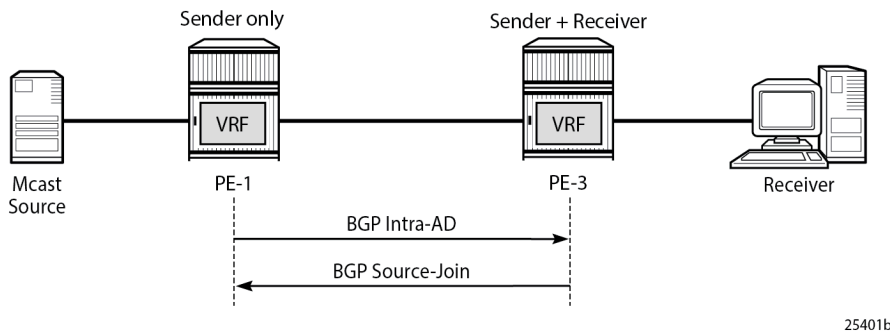
```
*A:PE-2# show router bgp neighbor 192.0.2.1 received-routes mvpn-ipv4
=====
BGP Router ID:192.0.2.2      AS:64500      Local AS:64500
=====
Legend -
Status codes : u - used, s - suppressed, h - history, d - decayed, * - valid
               l - leaked, x - stale, > - best, b - backup, p - purge
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP MVPN-IPv4 Routes
=====
Flag RouteType      OriginatorIP      LocalPref  MED
      RD            SourceAS          Path-Id     IGP Cost
      Nexthop       SourceIP          Label
      As-Path       GroupIP
-----
u*>i Intra-Ad          192.0.2.1        100        0
      64500:101     -                 None        -
      192.0.2.1     -
      No As-Path    -
-----
Routes : 1
=====
```

PE-2 output for the received routes for the mvpn-ipv6 address family is as follows:

```
*A:PE-2# show router bgp neighbor 192.0.2.1 received-routes mvpn-ipv6
=====
BGP Router ID:192.0.2.2      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 MVPN-IPv6 Routes
=====
Flag  RouteType      OriginatorIP      LocalPref  MED
      RD            SourceAS          Path-Id     IGP Cost
      Nexthop       SourceIP          Label
      As-Path       GroupIP
-----
u*>i  Intra-Ad          192.0.2.1        100        0
      64500:101      -                None        -
      192.0.2.1     -
      No As-Path    -
-----
Routes : 1
=====
```

The PE router does not change the BGP behavior when the MVPN is configured with **mdt-type sender-only**. A schematic of the message flow is presented in [Figure 58: mLDP-Based BGP Message Flow Between PE-1 and PE-3](#).

Figure 58: mLDP-Based BGP Message Flow Between PE-1 and PE-3



To demonstrate the BGP message flow sequence, the following initialization steps are taken:

1. Bring down the VPRN service, PIM protocol in the VPRN and IGMP/MLD protocol. As a result, the state of all signaling protocols is cleared.
2. Bring up the VPRN service. BGP exchanges unicast routing information.
3. Bring up the IPv4 PIM protocol. BGP exchanges IPv4 multicast routing information to build the PMSI infrastructure.
4. Bring up IGMP and add a static IGMP join where it is applicable. BGP exchanges IPv4 multicast routing information to propagate the multicast traffic to the receiver.

5. Bring up the IPv6 PIM protocol. BGP exchanges IPv6 multicast routing information to build the PMSI infrastructure.
6. Bring up MLD and add a static MLD join where it is applicable. BGP exchanges IPv6 multicast routing information to propagate the multicast traffic to the receiver.

The following BGP debug output is taken from PE-3 and demonstrates the message flow between PE-1 and PE-3.

The PMSI part is present in debug messages sent by PE-1 (**sender-only**).

Step 1: Bring down the VPRN service and protocols to clear the state of all signaling protocols.

```
# on PE-3
configure
  service
    vprn 1
      shutdown
      pim shutdown
      pim ipv6-multicast-disable
      igmp shutdown
      mld shutdown
    exit all
```

Step 2: Enable the VPRN service on PE-3. PE-3 immediately receives Intra-AD messages from PE-1 because the remote VPRN service is already enabled for IPv4 and IPv6 multicast propagation. The PMSI attribute is present in both messages.

```
# on PE-3
configure service vprn 1 no shutdown
```

```
15 2023/11/09 13:57:36.056 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Received BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 91
  Flag: 0x90 Type: 14 Len: 23 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.1
    Type: Intra-AD Len: 12 RD: 64500:101 Orig: 192.0.2.1
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:64500:1
  Flag: 0xc0 Type: 22 Len: 22 PMSI:
    Tunnel-type LDP P2MP LSP (2)
    Flags: (0x0)[Type: None BM: 0 U: 0 Leaf: not required]
    MPLS Label 0
    Root-Node 192.0.2.1, LSP-ID 0x2001
"
```

```
13 2023/11/09 13:57:36.055 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Received BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 91
  Flag: 0x90 Type: 14 Len: 23 Multiprotocol Reachable NLRI:
```

```

    Address Family MVPN_IPV6
    NextHop len 4 NextHop 192.0.2.1
    Type: Intra-AD Len: 12 RD: 64500:101 Orig: 192.0.2.1
    Flag: 0x40 Type: 1 Len: 1 Origin: 0
    Flag: 0x40 Type: 2 Len: 0 AS Path:
    Flag: 0x80 Type: 4 Len: 4 MED: 0
    Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
    Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
    Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:64500:1
    Flag: 0xc0 Type: 22 Len: 22 PMSI:
    Tunnel-type LDP P2MP LSP (2)
    Flags: (0x0)[Type: None BM: 0 U: 0 Leaf: not required]
    MPLS Label 0
    Root-Node 192.0.2.1, LSP-ID 0x2001
    "

```

Step 3: Enable PIM IPv4 only for the service on PE-3. PE-3 immediately sends Intra-AD messages to PE-1.

```

# on PE-3
configure service vprn 1 pim no shutdown

```

```

6 2023/11/09 13:57:34.867 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 91
  Flag: 0x90 Type: 14 Len: 23 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.3
    Type: Intra-AD Len: 12 RD: 64500:103 Orig: 192.0.2.3
    Flag: 0x40 Type: 1 Len: 1 Origin: 0
    Flag: 0x40 Type: 2 Len: 0 AS Path:
    Flag: 0x80 Type: 4 Len: 4 MED: 0
    Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
    Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
    Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:64500:1
    Flag: 0xc0 Type: 22 Len: 22 PMSI:
    Tunnel-type LDP P2MP LSP (2)
    Flags: (0x0)[Type: None BM: 0 U: 0 Leaf: not required]
    MPLS Label 0
    Root-Node 192.0.2.3, LSP-ID 0x2001
    "

```

Step 4: Bring up IGMP and add a static IGMP join for the service on a PE-3. PE-3 immediately sends a source-join message to PE-1 and receives a source-AD message from PE-1.

```

# on PE-3
configure service vprn 1 igmp no shutdown

```

```

# on PE-3
configure service vprn 1 igmp interface "int-PE-3-H-3"
static group 232.0.0.1 source 172.16.1.2

```

```

18 2023/11/09 13:57:40.877 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1

```

```
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 76
  Flag: 0x90 Type: 14 Len: 33 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.3
    Type: Source-Join Len:22 RD: 64500:101 SrcAS: 64500 Src: 172.16.1.2 Grp: 232.0.0.1
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:192.0.2.1:2
"
```

Step 5: Enable PIM IPv6 for the service on PE-3. PE-3 immediately sends Intra-AD messages to PE-1.

```
# on PE-3
configure service vprn 1 pim no ipv6-multicast-disable
```

```
20 2023/11/09 13:57:45.637 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 91
  Flag: 0x90 Type: 14 Len: 23 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV6
    NextHop len 4 NextHop 192.0.2.3
    Type: Intra-AD Len: 12 RD: 64500:103 Orig: 192.0.2.3
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:64500:1
  Flag: 0xc0 Type: 22 Len: 22 PMSI:
    Tunnel-type LDP P2MP LSP (2)
    Flags: (0x0)[Type: None BM: 0 U: 0 Leaf: not required]
    MPLS Label 0
    Root-Node 192.0.2.3, LSP-ID 0x2001
"
```

Step 6: Bring up MLD and add a static MLD join for the service on a PE-3. PE-3 immediately sends a source-join message to PE-1.

```
# on PE-3
configure service vprn 1 mld no shutdown
```

```
# on PE-3
configure service vprn 1 mld interface "int-PE-3-H-3" static
group ff3e::8000:1 source 2001:db8:1::2
```

```
22 2023/11/09 13:57:51.646 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.1
"Peer 1: 192.0.2.1: UPDATE
Peer 1: 192.0.2.1 - Send BGP UPDATE:
```

```

Withdrawn Length = 0
Total Path Attr Length = 100
Flag: 0x90 Type: 14 Len: 57 Multiprotocol Reachable NLRI:
  Address Family MVPN_IPV6
  NextHop len 4 NextHop 192.0.2.3
  Type: Source-Join Len:46 RD: 64500:101 SrcAS: 64500 Src: 2001:db8:1::2 Grp:
ff3e::8000:1
Flag: 0x40 Type: 1 Len: 1 Origin: 0
Flag: 0x40 Type: 2 Len: 0 AS Path:
Flag: 0x80 Type: 4 Len: 4 MED: 0
Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
Flag: 0xc0 Type: 8 Len: 4 Community:
  no-export
Flag: 0xc0 Type: 16 Len: 8 Extended Community:
  target:192.0.2.1:2
"
    
```

The same information can be gathered using the following show commands.

**show router bgp neighbor <peer> advertised-routes [mvpn-ipv4 | mvpn-ipv6]**

**show router bgp neighbor <peer> received-routes [mvpn-ipv4 | mvpn-ipv6]**

PE-3 output for the advertised routes for the mvpn-ipv4 address family is as follows:

```

*A:PE-3# show router bgp neighbor 192.0.2.1 advertised-routes mvpn-ipv4
=====
BGP Router ID:192.0.2.3      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 MVPN-IPv4 Routes
=====
Flag RouteType      OriginatorIP      LocalPref  MED
      RD            SourceAS          Path-Id     IGP Cost
      Nexthop       SourceIP          Label
      As-Path       GroupIP
-----
i     Source-Join    -                 100        0
      64500:101      64500            None       -
      192.0.2.3     172.16.1.2
      No As-Path    232.0.0.1
i     Intra-Ad      192.0.2.3        100        0
      64500:103     -                 None       -
      192.0.2.3     -
      No As-Path    -
-----
Routes : 2
=====
    
```

PE-3 output for the advertised routes for the mvpn-ipv6 address family is as follows:

```

*A:PE-3# show router bgp neighbor 192.0.2.1 advertised-routes mvpn-ipv6
=====
BGP Router ID:192.0.2.3      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 MVPN-IPv6 Routes
=====
Flag RouteType OriginatorIP LocalPref MED
RD SourceAS Path-Id IGP Cost
Nexthop SourceIP Label
As-Path GroupIP
-----
i Source-Join - 100 0
64500:101 64500 None -
192.0.2.3 2001:db8:1::2
No As-Path ff3e::8000:1
i Intra-Ad 192.0.2.3 100 0
64500:103 - None -
192.0.2.3 -
No As-Path -
-----
Routes : 2
=====

```

PE-3 output for the received routes for the mvpn-ipv4 address family is as follows:

```

*A:PE-3# show router bgp neighbor 192.0.2.1 received-routes mvpn-ipv4
=====
BGP Router ID:192.0.2.3 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 MVPN-IPv4 Routes
=====
Flag RouteType OriginatorIP LocalPref MED
RD SourceAS Path-Id IGP Cost
Nexthop SourceIP Label
As-Path GroupIP
-----
u*>i Intra-Ad 192.0.2.1 100 0
64500:101 - None -
192.0.2.1 -
No As-Path -
-----
Routes : 1
=====

```

PE-3 output for the received routes for the mvpn-ipv6 address family is as follows:

```

*A:PE-3# show router bgp neighbor 192.0.2.1 received-routes mvpn-ipv6
=====
BGP Router ID:192.0.2.3 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 MVPN-IPv6 Routes
=====

```

Flag	RouteType	OriginatorIP	LocalPref	MED
	RD	SourceAS	Path-Id	IGP Cost
	Nexthop	SourceIP		Label
	As-Path	GroupIP		
u*>i	Intra-Ad	192.0.2.1	100	0
	64500:101	-	None	-
	192.0.2.1	-		
	No As-Path	-		
-----				
Routes : 1				
=====				

## LDP Verification and Debugging

When BGP intra-AD messages are exchanged, every PE starts to build a multicast tunnel based on the following criteria:

- PE nodes which are configured as **sender-only** do not distribute mLDP forward equivalence classes (FECs) to remote PEs for this MVPN.
- PE nodes which are configured as **receiver-only** do not include the PMSI part for intra-AD messages and remote PEs do not send mLDP FECs for this MVPN.

LDP bindings can be verified using the following command:

### show router ldp bindings p2mp

PE-1 (192.0.2.1) has two egress FECs because PE-1 has **mdt-type sender-only**.

```
*A:PE-1# show router ldp bindings p2mp ipv4
=====
LDP Bindings (IPv4 LSR ID 192.0.2.1)
              (IPv6 LSR ID 2001:db8::1)
=====
Label Status:
  U - Label In Use, N - Label Not In Use, W - Label Withdrawn
  WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
  e - Label ELC
FEC Flags:
  LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
  BA - ASBR Backup FEC
=====
LDP Generic IPv4 P2MP Bindings
=====
P2MP-Id
RootAddr          Interface
Peer
IngLbl            EgrLbl
EgrNH             EgrIf/LspId
-----
8193
192.0.2.1         73729
192.0.2.2:0      524279
--
192.168.12.2     1/1/c1/1

8193
192.0.2.1         73729
192.0.2.3:0      524280
--
```

```

192.168.13.2                               1/1/c2/1
-----
No. of Generic IPv4 P2MP Bindings: 2
=====
---snip---
=====

```

PE-2 (192.0.2.2) has two ingress FECs because PE-2 has **mdt-type receiver-only**.

```

*A:PE-2# show router ldp bindings p2mp ipv4
=====
LDP Bindings (IPv4 LSR ID 192.0.2.2)
              (IPv6 LSR ID 2001:db8::2)
=====
Label Status:
  U - Label In Use, N - Label Not In Use, W - Label Withdrawn
  WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
  e - Label ELC
FEC Flags:
  LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
  BA - ASBR Backup FEC
=====
LDP Generic IPv4 P2MP Bindings
=====
P2MP-Id      RootAddr      Interface
Peer
IngLbl       EgrLbl
EgrNH        EgrIf/LspId
-----
8193
192.0.2.1    73733
192.0.2.1:0
524279U     --
--         --
8193
192.0.2.3    73734
192.0.2.3:0
524280U     --
--         --
-----
No. of Generic IPv4 P2MP Bindings: 2
=====
---snip---
=====

```

PE-3 (192.0.2.3) has one ingress FEC and one egress FECs because PE-3 has the default **mdt-type sender-receiver**. There is only an egress FEC to PE-2 (receiver-only), but not to PE-1. PE-1 can never be a receiver, because it is configured as sender-only.

```

*A:PE-3# show router ldp bindings p2mp ipv4 opaque-type generic
=====
LDP Bindings (IPv4 LSR ID 192.0.2.3)
              (IPv6 LSR ID 2001:db8::3)
=====
Label Status:
  U - Label In Use, N - Label Not In Use, W - Label Withdrawn

```

```

WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
e - Label ELC
FEC Flags:
LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
BA - ASBR Backup FEC
=====
LDP Generic IPv4 P2MP Bindings
=====
P2MP-Id
RootAddr                Interface
Peer
IngLbl                  EgrLbl
EgrNH                   EgrIf/LspId
-----
8193
192.0.2.1                73733
192.0.2.1:0
524280U                  --
--                        --

8193
192.0.2.3                73732
192.0.2.2:0
--                        524280
192.168.23.1            1/1/c2/1
-----
No. of Generic IPv4 P2MP Bindings: 2
=====

```

## Multicast Stream Verification

The status of a multicast group/stream can be verified using the following command:

**show router <sid> pim group detail [ipv6]**

An IPv4 sender is connected to PE-1. The physical interface where the source is connected is used as incoming interface and the I-PMSI is used as outgoing interface.

```

*A:PE-1# show router 1 pim group detail
=====
PIM Source Group ipv4
=====
Group Address       : 232.0.0.1
Source Address      : 172.16.1.2
RP Address          : 0
Advt Router        : 192.0.2.1
Flags              :
Mode               : sparse
MRIB Next Hop      : 172.16.1.2
MRIB Src Flags     : direct
Keepalive Timer    : Not Running
Up Time            : 0d 00:05:30
Resolved By        : rtable-u

Up JP State        : Joined
Up JP Rpt          : Not Joined StarG
Up JP Expiry       : 0d 00:00:00
Up JP Rpt Override : 0d 00:00:00

Register State     : No Info
Reg From Anycast RP: No

```

```

Rpf Neighbor      : 172.16.1.2
Incoming Intf    : int-PE-1-S-1
Outgoing Intf List : mpls-if-73729

Curr Fwding Rate : 4811.800 kbps
Forwarded Packets : 202316          Discarded Packets : 0
Forwarded Octets  : 198674312      RPF Mismatches    : 0
Spt threshold     : 0 kbps          ECMP opt threshold : 7
Admin bandwidth   : 1 kbps
-----
Groups : 1
=====

```

There is an IPv4 receiver connected to PE-2. The I-PMSI is used as incoming interface and the physical interface where the receiver is connected is used as outgoing.

```

*A:PE-2# show router 1 pim group detail

=====
PIM Source Group ipv4
=====
Group Address      : 232.0.0.1
Source Address     : 172.16.1.2
RP Address         : 0
Advt Router        : 192.0.2.1
Flags              :                               Type           : (S,G)
Mode               : sparse
MRIB Next Hop      : 192.0.2.1
MRIB Src Flags     : remote
Keepalive Timer    : Not Running
Up Time            : 0d 00:03:05      Resolved By         : rtable-u

Up JP State        : Joined           Up JP Expiry        : 0d 00:00:54
Up JP Rpt          : Not Joined StarG Up JP Rpt Override  : 0d 00:00:00

Register State     : No Info
Reg From Anycast RP: No

Rpf Neighbor      : 192.0.2.1
Incoming Intf     : mpls-if-73733
Outgoing Intf List : int-PE-2-H-2

Curr Fwding Rate   : 4811.800 kbps
Forwarded Packets  : 113493          Discarded Packets   : 0
Forwarded Octets   : 111450126      RPF Mismatches     : 0
Spt threshold      : 0 kbps          ECMP opt threshold  : 7
Admin bandwidth    : 1 kbps
-----
Groups : 1
=====

```

There is IPv4 receiver connected to PE-3. The I-PMSI is used as incoming interface and the physical interface where the receiver is connected is used as outgoing.

```

*A:PE-3# show router 1 pim group detail

=====
PIM Source Group ipv4
=====
Group Address      : 232.0.0.1
Source Address     : 172.16.1.2
RP Address         : 0

```

```

Advt Router      : 192.0.2.1
Flags           :                               Type           : (S,G)
Mode           : sparse
MRIB Next Hop   : 192.0.2.1
MRIB Src Flags  : remote
Keepalive Timer : Not Running
Up Time        : 0d 00:01:04      Resolved By       : rtable-u

Up JP State     : Joined           Up JP Expiry      : 0d 00:00:55
Up JP Rpt      : Not Joined StarG Up JP Rpt Override : 0d 00:00:00

Register State  : No Info
Reg From Anycast RP: No

Rpf Neighbor    : 192.0.2.1
Incoming Intf : mpls-if-73733
Outgoing Intf List : int-PE-3-H-3

Curr Fwding Rate : 4811.800 kbps
Forwarded Packets : 39761          Discarded Packets : 0
Forwarded Octets  : 39045302      RPF Mismatches    : 0
Spt threshold    : 0 kbps          ECMP opt threshold : 7
Admin bandwidth  : 1 kbps
-----
Groups : 1
=====

```

Similar behavior is observed for IPv6 multicast.

There is an IPv6 sender connected to PE-1. The physical interface where the sender is connected is used as the incoming interface and the I-PMSI is used as the outgoing interface.

```

*A:PE-1# show router 1 pim group detail ipv6
=====
PIM Source Group ipv6
=====
Group Address      : ff3e::8000:1
Source Address     : 2001:db8:1::2
RP Address         : 0
Advt Router       : 192.0.2.1
Flags             :                               Type           : (S,G)
Mode             : sparse
MRIB Next Hop     : 2001:db8:1::2
MRIB Src Flags    : direct
Keepalive Timer   : Not Running
Up Time          : 0d 00:05:30      Resolved By       : rtable6-u

Up JP State       : Joined           Up JP Expiry      : 0d 00:00:00
Up JP Rpt        : Not Joined StarG Up JP Rpt Override : 0d 00:00:00

Register State    : No Info
Reg From Anycast RP: No

Rpf Neighbor      : 2001:db8:1::2
Incoming Intf   : int-PE-1-S-1
Outgoing Intf List : mpls-if-73729

Curr Fwding Rate  : 4815.728 kbps
Forwarded Packets : 202316          Discarded Packets : 0
Forwarded Octets  : 198674312      RPF Mismatches    : 0
Spt threshold    : 0 kbps          ECMP opt threshold : 7
Admin bandwidth  : 1 kbps

```

```
-----
Groups : 1
=====
```

There is an IPv6 receiver connected to PE-2. An I-PMSI is used as the incoming interface and the physical interface where the receiver is connected is used as the outgoing interface.

```
*A:PE-2# show router 1 pim group detail ipv6

=====
PIM Source Group ipv6
=====
Group Address      : ff3e::8000:1
Source Address     : 2001:db8:1::2
RP Address         : 0
Advt Router        : 192.0.2.1
Flags              :                               Type           : (S,G)
Mode               : sparse
MRIB Next Hop      : 192.0.2.1
MRIB Src Flags     : remote
Keepalive Timer    : Not Running
Up Time            : 0d 00:02:51      Resolved By         : rtable6-u

Up JP State        : Joined           Up JP Expiry        : 0d 00:00:08
Up JP Rpt          : Not Joined StarG Up JP Rpt Override  : 0d 00:00:00

Register State     : No Info
Reg From Anycast RP: No

Rpf Neighbor       : 192.0.2.1
Incoming Intf    : mpls-if-73733
Outgoing Intf List : int-PE-2-H-2

Curr Fwding Rate   : 4815.728 kbps
Forwarded Packets  : 104785           Discarded Packets   : 0
Forwarded Octets   : 102898870       RPF Mismatches      : 0
Spt threshold      : 0 kbps           ECMP opt threshold  : 7
Admin bandwidth    : 1 kbps

-----
Groups : 1
=====
```

There is an IPv6 receiver connected to PE-3. An I-PMSI is used as the incoming interface and the physical interface where the receiver is connected is used as the outgoing interface.

```
*A:PE-3# show router 1 pim group detail ipv6

=====
PIM Source Group ipv6
=====
Group Address      : ff3e::8000:1
Source Address     : 2001:db8:1::2
RP Address         : 0
Advt Router        : 192.0.2.1
Flags              :                               Type           : (S,G)
Mode               : sparse
MRIB Next Hop      : 192.0.2.1
MRIB Src Flags     : remote
Keepalive Timer    : Not Running
Up Time            : 0d 00:00:54      Resolved By         : rtable6-u

Up JP State        : Joined           Up JP Expiry        : 0d 00:00:06
```

```
Up JP Rpt          : Not Joined StarG   Up JP Rpt Override : 0d 00:00:00

Register State     : No Info
Reg From Anycast RP: No

Rpf Neighbor       : 192.0.2.1
Incoming Intf    : mpls-if-73733
Outgoing Intf List : int-PE-3-H-3

Curr Fwding Rate   : 4811.800 kbps
Forwarded Packets  : 33162
Forwarded Octets   : 32565084
Spt threshold      : 0 kbps
Admin bandwidth    : 1 kbps

Discarded Packets  : 0
RPF Mismatches    : 0
ECMP opt threshold : 7

-----
Groups : 1
=====
```

## Conclusion

The sender-only/receiver-only feature provides significant signaling optimization in the core network for RSVP and LDP protocols and is recommended to be used when such functionality is required. The following are required before implementing this feature in the network:

- MDT-types **sender-only**, **receiver-only** and **sender-receiver** are enabled per MVPN.
- The default mdt-type is **sender-receiver** mode for backward compatibility.
- This is purely a control plane feature and there are no hardware dependencies.
- Rosen MPVN or MDT-SAFI based MVPNs are not supported.
- IPv4 and IPv6 C-signaling are supported.
- mLDP and RSVP demonstrate slightly different behavior because of the nature of each protocol.
- mLDP provides a better optimization than RSVP in all cases, as mLDP does not initiate LSPs to sender-only routers.



---

# NG-MVPN Source Redundancy

This chapter provides information about MVPN source redundancy.

Topics in this chapter include:

- [Applicability](#)
- [Overview](#)
- [Configuration](#)
- [Conclusion](#)

## Applicability

The chapter was initially written for SR OS Release 12.0.R1, using multicast LDP as the provider tunnel signaling mechanism for IPv4 multicasting. The customer multicast signaling protocol within the VPN must be BGP. The CLI in the current edition corresponds with SR OS Release 23.10.R2.

## Overview

Multicast source redundancy allows operators to provide multiple geo-redundant sources for the same multicast group in a multicast virtual private network (MVPN). For instance, in an IPTV environment where a TV channel maps to a multicast group, the same TV channel can be provided from sources in a geographically diverse manner where a national broadcaster can have multiple sources from two or more regional distribution centers.

Knowledge of Multi-Protocol BGP (MP-BGP) and RFC 4364, *BGP/MPLS IP Virtual Private Networks (VPNs)*, is assumed throughout this chapter, as well as Protocol Independent Multicast (PIM), RFC 6513, *Multicast in MPLS/BGP IP VPNs*, and RFC 6514, *BGP Encodings and Procedures for Multicast in MPLS/BGP IP VPNs*.

Hosts connected to receiver PEs can receive TV channels from a specific source, with a regional backup source available in case of a failure.

Figure 59: Source Redundancy Example.

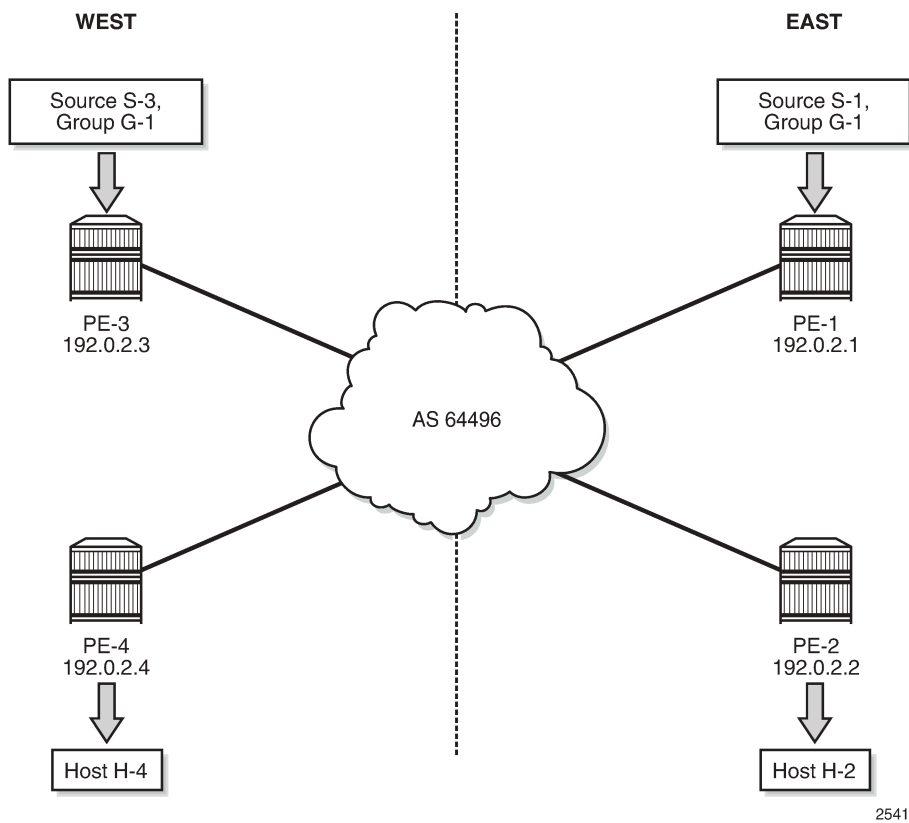


Figure 59: Source Redundancy Example. shows the concept of source redundancy. PE-1 and PE-3 have directly connected multicast sources. For clarity, consider a single multicast group with two separate sources connected at different sites. The content of each group is identical at a given time (allowing for transmission delay), as is expected for an IPTV channel. PE-1 and PE-3 are referred to as sender PEs because they are closer to the source; PE-2 and PE-4 are referred to as receiver PEs because they are closer to hosts H-2 and H-4.

A multicast group, group 1 (G-1) has two sources: Source 1 (S-1) in the east region and Source 3 (S-3) in the west region which are connected to PE-1 and PE-3 respectively. Receivers connected to PE-2 in the east region join group (S-1,G-1) and receivers connected to PE-4 in the west region join group (S-3,G-1). The presence of each source is declared within the multicast VPN by the sender PE. When a multicast group becomes active, a BGP Source Active auto-discovery (SA) route is advertised to all PEs within the multicast VPN. This must occur even if no receiver indicates that it wants to become a member of this group. In other words, the SA must be persistently present in the receiver PEs when the source is available.

Should either source fail or become unavailable, then the sender PE notifies the receiver PEs by sending an NLRI unreachable BGP SA Route that declares the absence of the source. All hosts that are members of this group then switch to receive traffic from the remaining active source. Only customer multicast joins received as IGMP (\*,G) queries or PIM (\*,G) joins at the receiver PE are valid, because the source address is not specified.

Source redundancy is achieved by:

- Configuring a list of redundant sources within each receiver PE.
- Configuring the sender PEs to originate a BGP Source Active Auto Discovery for each detected active multicast source, regardless of whether a receiver is joined to the multicast group or not. As a result, a Source Active route is originated on a per (S,G) basis.

For multiple SAs to be persistently present in the receiver PEs, one of the following two conditions must be configured within the sender PEs:

- Either disable inter-site shared trees on the sender PEs, such that there is no c-tree with root at the RP. Any active source announces its presence using a BGP SA to all receiver PEs so no shared joins are sent by receiver PEs to RP, or
- Leave inter-site shared trees as enabled, but configured so that the SA AD route for each multicast group is persistently present in the receiver PEs, even in the absence of requesting hosts for each group. Shared and Source Joins are sent by the receiver PEs.

Both of the preceding options are supported. The default behavior has inter-site shared trees enabled without persistency. In this example, inter-site shared trees at the sender PEs are enabled with Source Active routes set to be persistent.

- Ensuring that the preferred source is IP reachable within the VPRN from the receiver PE. This must be a remote source advertised from a remote PE within the VPRN.
- Receiver PEs accept the Source Active routes into the appropriate Multicast VRF.
- Ensuring the preferred active source should have a higher BGP Local Preference. This is achieved using a route policy. Any other sources from the redundant list should exist as suppressed standby sources, but the (S,G) state should exist if the source is active – when a valid BGP MVPN Source Active route for that source has been received.

All of these conditions are achieved by configuration.

To allow each receiver PE to choose a preferred source, each SA route advertised by the sender PE is tagged with a community value. Each receiver PE can then use the community value contained within each SA route update received to set the Local Preference BGP attribute to a value such that the receiver PE can choose the most preferred active source.

The objectives are:

- To configure multicast in a VPRN on PE-1 to PE-4 with inter-site-shared trees enabled on the receiver PEs and Source Active routes persistently present, for reasons previously described.
- To connect redundant sources to the sender PE-1 and PE-3, with each multicast source having the same group address. For ease of configuration, a single redundant source is used.
- To advertise each source to the receiver PEs (PE-2 and PE-4), using appropriate route policies for adding community strings to the BGP Source Active Auto-Discovery routes.
- To configure appropriate route policies that allow each BGP SA route to have the correct Local Preference set, based on the community strings present.
- To allow receivers to connect to the appropriate source, using (\*,G) joins.

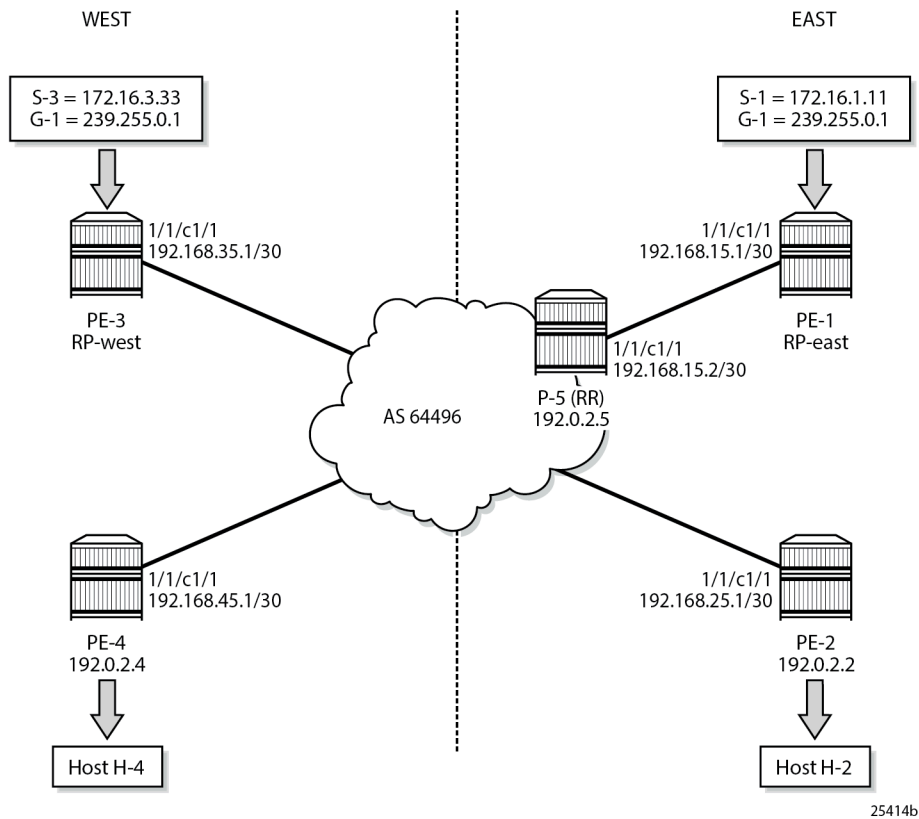
The following configuration tasks should be completed as a pre-requisite:

- Full mesh IS-IS or OSPF between each of the PE routers and the route reflector.
- Link-layer LDP between all PEs. RSVP could also be used.
- Multicast LDP is used as the provider tunnel signaling protocol. This is enabled by default when link layer LDP is enabled. RSVP and PIM SSM are also supported as provider tunnel signaling mechanisms and could be used.

## Configuration

The example topology is shown in [Figure 60: Schematic Topology](#), containing the four PEs plus the route reflector on P-5.

Figure 60: Schematic Topology



## Global BGP Configuration

The first step is to configure an iBGP session between each of the PEs and the route reflector (RR) seen in [Figure 60: Schematic Topology](#). The address families negotiated between the iBGP peers are `vpn-ipv4` (unicast routing) and `mvpn-ipv4` (multicast routing). The BGP configuration for all PE nodes is identical:

```
# on PE-1
configure
router
  bgp
    rapid-update mvpn-ipv4
    group "INTERNAL"
      family vpn-ipv4 mvpn-ipv4
      type internal
      neighbor 192.0.2.5
    exit
  exit
```

```
no shutdown
exit all
```

The configuration for the Route Reflector on P-5 is as follows:

```
# on P-5
configure
router
  bgp
    rapid-update mvpn-ipv4
    group "RRclients"
      family vpn-ipv4 mvpn-ipv4
      type internal
      cluster 1.1.1.1
      neighbor 192.0.2.1
      exit
      neighbor 192.0.2.2
      exit
      neighbor 192.0.2.3
      exit
      neighbor 192.0.2.4
      exit
    exit
  no shutdown
exit all
```

On PE-1, verify that the BGP session with RR on P-5 is established with address families "vpn-ipv4" and "mvpn-ipv4" capabilities negotiated:

```
*A:PE-1# show router bgp summary
=====
BGP Router ID:192.0.2.1      AS:64496      Local AS:64496
=====
BGP Admin State      : Up      BGP Oper State      : Up
Total Peer Groups    : 1        Total Peers          : 1
Total VPN Peer Groups : 0        Total VPN Peers      : 0
Current Internal Groups : 1      Max Internal Groups  : 1
Total BGP Paths       : 15       Total Path Memory    : 5520

---snip---
=====
BGP Summary
=====
Legend : D - Dynamic Neighbor
=====
Neighbor
Description
          AS PktRcvd InQ  Up/Down  State|Rcv/Act/Sent (Addr Family)
          PktSent OutQ
-----
192.0.2.5
          64496      5    0 00h00m40s 0/0/0 (VpnIPv4)
          5         0    0/0/0 (MvpnIPv4)
-----
```

The same command can be used on the other PEs to verify their BGP sessions to the RR.

## Configuring VPRN on PEs

The following outputs show the VPRN configurations for each PE. The specific MVPN configuration is shown later.

### PE-1

The VPRN configuration for PE-1 is as follows:

```
# on PE-1
configure
service
  vprn 1 customer 1 create
    interface "int-PE-1-S-1" create
      address 172.16.1.1/24
      sap 1/1/c3/1 create
      exit
    exit
    interface "RP" create
      address 10.10.10.1/32
      loopback
    exit
    bgp-ipvpn
      mpls
        auto-bind-tunnel
        resolution-filter
        ldp
        exit
        resolution filter
      exit
      route-distinguisher 64496:1
      vrf-target target:64496:1
      no shutdown
    exit
  exit
  pim
    apply-to all
    rp
      static
        address 10.10.10.1
        group-prefix 239.0.0.0/8
      exit
    exit
  exit
  no shutdown
  exit
exit all
```

There is a single interface toward S-1 from which the multicast group is generated.

If the customer signaling uses PIM ASM, then the customer Rendezvous Point (RP) must be positioned on the sender PE because registration of the source with the RP causes the SA to be sent to the remote source PEs.

A loopback interface called "RP" acts as the RP for all group prefixes in the 239.0.0.0/8 range. This is the RP for the East groups.

### PE-2

PE-2 has a receiver attached, and a single interface is configured to accommodate this. The RP configured is that of the East region and has a configuration as follows:

```
# on PE-2
configure
service
  vprn 1 customer 1 create
  interface "int-PE-2-H-2" create
  address 172.16.2.1/24
  sap 1/1/c3/1 create
  exit
exit
bgp-ipvpn
  mpls
  auto-bind-tunnel
  resolution-filter
  ldp
  exit
  resolution filter
  exit
  route-distinguisher 64496:1
  vrf-target target:64496:1
  no shutdown
  exit
exit
igmp
  interface "int-PE-2-H-2"
  version 2 # when IGMPv2 is used
  exit
exit
pim
  apply-to all
  rp
  static
  address 10.10.10.1
  group-prefix 239.0.0.0/8
  exit
  exit
  exit
  no shutdown
  exit
exit all
```

### PE-3

PE-3 serves as the RP for the West region and uses a different IP address for the Rendezvous Point interface.

```
# on PE-3
configure
service
  vprn 1 customer 1 create
  interface "int-PE-3-S-3" create
  address 172.16.3.1/24
  sap 1/1/c3/1 create
  exit
exit
interface "RP" create
  address 10.10.10.3/32
  loopback
  exit
bgp-ipvpn
  mpls
```

```

        auto-bind-tunnel
        resolution-filter
        ldp
        exit
        resolution filter
    exit
    route-distinguisher 64496:1
    vrf-target target:64496:1
    no shutdown
exit
exit
pim
    apply-to all
    rp
        static
        address 10.10.10.3
        group-prefix 239.0.0.0/8
        exit
    exit
exit
no shutdown
exit
exit all

```

#### PE-4

PE-4 also has a receiver, and uses the West sender PE (PE-3) as the Rendezvous Point.

```

# on PE-4
configure
service
    vprn 1 customer 1 create
    interface "int-PE-4-H-4" create
        address 172.16.4.1/24
        sap 1/1/c3/1 create
        exit
    exit
    bgp-ipvpn
    mpls
        auto-bind-tunnel
        resolution-filter
        ldp
        exit
        resolution filter
    exit
    route-distinguisher 64496:1
    vrf-target target:64496:1
    no shutdown
    exit
exit
igmp
    interface "int-PE-4-H-4"
        version 2 # when IGMPv2 is used
    exit
exit
pim
    apply-to all
    rp
        static
        address 10.10.10.3
        group-prefix 239.0.0.0/8
        exit
    exit
exit
exit

```



```

no shutdown
exit
exit all

```

## MVPN Configuration for Source PEs

At the PEs closest to the sources (PE-1 and PE-3), Source Active auto-discovery BGP routes are generated when the source is active.

This applies for PIM-ASM (\*,G) joins only, or IGMP (\*,G) membership queries received by the provider domain. These are received by all PEs.

Inter-site trees must be disabled for this to occur. Alternatively, inter-site trees can be enabled such that when a source is discovered, a Source Active is advertised to each other PE in the MVPN. This occurs regardless of whether any receivers want to become members of the multicast groups.

As previously stated, the presence of the SA in the receiver PEs means that no shared joins routes are generated toward the C-RPs.

The MVPN configuration enables BGP as both auto-discovery mechanism and the customer multicast signaling protocol across the VPRN. The provider tunnel between PEs within the MVPN is signaled using Multicast LDP.

The MVPN configuration for each PE should be as follows:

```

# on PE-1
configure
service
vprn 1
mvpn
    auto-discovery default
    c-mcast-signaling bgp
    provider-tunnel
        inclusive
        mldp
        no shutdown
    exit
exit
exit
exit
no shutdown
exit all

```

The VPRN MVPN configuration for PE-2, PE-3, and PE-4 is identical.

## Sender PE Route Policies

The choice of active and standby sources by the receiver PEs is determined by the “best route” policy. PE-1 and PE-3 advertise BGP Source Active Auto Discovery routes when a source is active. This is received by all PEs within the MVPN. As two different sources advertise the same group, it is necessary to differentiate between them.

Assuming that receiver PE-2 prefers the source from PE-1, and PE-4 prefers the source active on PE-3, then the export policy for MVPN routes on PE-1 requires the following steps:

1. Set a community value on PE-1 for the (S,G) multicast group – call this “blue” with value 64496:11.
2. Set the route target community for the VPRN – 64496:1.

3. Create a policy statement that becomes the export policy for MVPN routes within PE-1.
4. Create a policy statement entry (entry 10) that adds the community value "blue" along with the route target for Source Active AD BGP routes. Source Active AD routes are MVPN type 5 routes.
5. Create a policy statement entry default-action that adds the route target for all other MVPN AD BGP routes (such as Intra-AD (type 1)) that are exported to the MVPN PEs.

```
# on PE-1
configure
router
  policy-options
  begin
    community "blue" members "64496:11"
    community "MVPN1_RT" members "target:64496:1"
    policy-statement "MVPN1_export"
    entry 10
      description "match MVPN routes - type 5 Source AD - add RT and
'blue' community"
      from
        mvpn-type 5
        family mvpn-ipv4
      exit
      action accept
        community add "blue" "MVPN1_RT"
      exit
    exit
    default-action accept
      community add "MVPN1_RT"
    exit
  exit
  commit
exit all
```

6. Apply as an export policy within the MVPN context.

The import policy requires that all imported MVPN BGP routes have the correct route target extended community value, specifically "target:64496:1".

1. Create a policy statement that becomes the import policy for PE-1.
2. Create a policy statement entry (entry 10) that matches the community of the route target extended community for all MVPN BGP routes. These include the Intra-AD and Source-Join routes.

```
# on PE-1
configure
router
  policy-options
  begin
    policy-statement "MVPN1_import"
    entry 10
      from
        community "MVPN1_RT"
      exit
      action accept
      exit
    exit
  exit
  commit
exit all
```

Enable the inter-site-shared type 5 advertisement persistency so that source ADs are advertised when multicast sources are active. Alternatively, inter-site shared trees can be disabled using the **no intersite-shared** command. In this example, only inter-site shared MVPN type 5 persistency is shown.

The additional configuration in the MVPN context is as follows, where the PIM instance must be shut down when the intersite-shared configuration is modified.

```
# on PE-1
configure
  service
    vprn 1
      pim shutdown      # needed to modify mvpn configuration
      mvpn
        intersite-shared persistent-type5-adv
        vrf-import "MVPN1_import"
        vrf-export "MVPN1_export"
        mdt-type sender-only  # added to prevent a sender PE from also behaving as a
receiver PE
      exit
      pim no shutdown
    exit all
```

For PE-3 (the other sender PE), similar import and export policies are required. In this case, the community is called "red" and is added to the Source Active AD route generated when the source is active.

The requirements for the export policy for PE-3 are as follows:

```
# on PE-3
configure
  router
    policy-options
      begin
        community "red" members "64496:33"
        community "MVPN1_RT" members "target:64496:1"
        policy-statement "MVPN1_export"
          entry 10
            description "match MVPN routes - type 5 Source AD - add RT and 'red'"
community"
            from
              mvpn-type 5
              family mvpn-ipv4
            exit
            action accept
              community add "red" "MVPN1_RT"
            exit
          exit
          default-action accept
            community add "MVPN1_RT"
          exit
        exit
      exit
    commit
  exit all
```

The import policy is exactly the same as for PE-1.

Apply the import and export policies to the MVPN context of the sender PE (PE-3) and enable inter-site-shared type 5 advertisement persistency with the same command as on PE-1.

## Receiver PE Configuration

PE-2 and PE-4 are the receiver PEs. These receive the Source Active AD routes and initiate Joins toward the preferred source.

When a Source-Active AD route is received, the community value is examined and the Local Preference value of the route is set using a Route Policy. The preferred source is determined by the SA AD route with the highest Local Preference value.

In the case of PE-2, the preferred source is that advertised by PE-1, the "blue" source as previously referenced. PE-2 sets the Local Preference to 200. The SA AD tagged with the "red" community has the Local Preference set to 50.

For PE-4, the reverse applies: SA AD routes tagged with the "red" community have the Local Preference set to 200, and "blue" SA AD routes have the Local Preference set to 50.

Again, assuming that the PE-2 receiver prefers the source from PE-1 and PE-4 prefers the source active on PE-3, the import policy for MVPN routes on PE-2 requires the following steps:

1. Set a community value on PE-2 for the (S,G), call this "blue" with value 64496:11.
2. Set the route target community for the VPRN to 64496:1.
3. Create a prefix list that matches the multicast group address, in this case 239.255.0.0/24.
4. Create a policy statement that becomes the import policy for MVPN routes within PE-1.
5. Create a policy statement entry (entry 10) that matches the following attributes:
  - Source Active AD BGP routes type. Source Active AD routes are classed as MVPN type 5 routes, and
  - Community value "blue" AND Route Target extended community, and
  - Group address prefix 239.255.0.0/24

If the BGP route matches all three conditions, then set the Local Preference to 200.

6. Create a policy statement default-action that accepts all other MVPN BGP routes, including SA routes tagged with the "red" community value.

The import policy statement looks like:

```
# on PE-2
configure
router
    policy-options
        begin
            prefix-list "group_239.255.x.y"
                prefix 239.255.0.0/16 longer
            exit
            community "blue" members "64496:11"
            community "red" members "64496:33"
            community "MVPN1_RT" members "target:64496:1"
            policy-statement "MVPN1_import"
                entry 10
                    description "allow MVPN source-ad - set LP to 200 for 'blue'"
                    from
                        community expression "[blue] AND [MVPN1_RT]"
                        mvpn-type 5
                        group-address "group_239.255.x.y"
                    exit
                    action accept
```

```

        local-preference 200
        exit
    exit
    entry 20
        description "allow MVPN source-ad - set LP to 50 for 'red'"
        from
            community expression "[red] AND [MVPN1_RT]"
            mvpn-type 5
            group-address "group_239.255.x.y"
        exit
        action accept
            local-preference 50
        exit
    exit
    default-action accept
    exit
    exit
    commit
    exit all

```

The export policy for PE-2 MVPN routes requires each MVPN route to be tagged with the route target extended community for VPRN 1. The following policy statement is created:

```

# on PE-2
configure
    router
        policy-options
            begin
                policy-statement "MVPN1_export"
                    entry 10
                        from
                            family mvpn-ipv4
                        exit
                        action accept
                            community add "MVPN1_RT"
                        exit
                    exit
                exit
            commit
        exit all

```

7. Create a list of redundant sources. This is a list of prefixes that match the source addresses of redundant multicast groups. This is an important parameter because the receiver PEs only create active and standby (S,G) states for groups with source address prefixes that are contained in this list.
8. Before any hosts attempt to join the multicast groups, the decision must be made to enable or disable inter-site shared trees at the receiver PEs. In this example, only the Inter-site shared trees disabled option is considered. To make this configuration change, it is necessary to shut the PIM protocol down before and re-enable when completed.

The additional MVPN configuration for PE-2 is shown in the following output, where the redundant source prefix list is included, and inter-site shared trees are disabled.

```

# on PE-2
configure
    service
        vprn 1
            pim shutdown # needed to modify mvpn configuration
            mvpn
                no intersite-shared
                red-source-list

```

```

        src-prefix 172.16.1.0/24
        src-prefix 172.16.3.0/24
    exit
    vrf-import "MVPN1_import"
    vrf-export "MVPN1_export"
    mdt-type receiver-only # added to prevent a receiver PE from also behaving
as a sender PE
    exit
    pim no shutdown
    exit all

```

PE-4 requires a similar set of import and export policies. In this case, the “red” sources have the highest Local Preference value, based on the community string added by the export policy of PE-3.

```

# on PE-4
configure
router
  policy-options
  begin
    prefix-list "group_239.255.x.y"
    prefix 239.255.0.0/16 longer
    exit
    community "blue" members "64496:11"
    community "red" members "64496:33"
    community "MVPN1_RT" members "target:64496:1"
    policy-statement "MVPN1_import"
    entry 10
      description "allow MVPN source-ad - set LP to 200 for 'red'"
      from
        community expression "[red] AND [MVPN1_RT]"
        mvpn-type 5
        group-address "group_239.255.x.y"
      exit
      action accept
      local-preference 200
    exit
    entry 20
      description "allow MVPN source-ad - set LP to 50 for 'blue'"
      from
        community expression "[blue] AND [MVPN1_RT]"
        mvpn-type 5
        group-address "group_239.255.x.y"
      exit
      action accept
      local-preference 50
    exit
    default-action accept
  exit
  commit
exit all

```

The export policy for MVPN routes adds the route target extended community. It is exactly the same export policy as for PE-2.

The additional MVPN configuration for VPRN 1 on PE-4 is identically the same as for PE-2.

Each PE within the MVPN originates an Intra-AD BGP route. This notifies the other PEs within the VPRN. This is used to create a set of Inclusive Provider Multicast Service Interfaces (I-PMSI) between each PE. In this case, I-PMSIs are signaled using mLDP.

Using PE-1 as an example, the set of Intra-AD routes can be seen using the following command:

```
*A:PE-1# show router bgp routes mvpn-ipv4
=====
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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP MVPN-IPv4 Routes
=====
Flag  RouteType      OriginatorIP      LocalPref  MED
      RD            SourceAS          Path-Id     IGP Cost
      Nexthop       SourceIP          Label
      As-Path       GroupIP
-----
i     Intra-Ad        192.0.2.1        100        0
      64496:1        -                None       -
      192.0.2.1    -                -
      No As-Path   -
u*>i Intra-Ad        192.0.2.2        100        0
      64496:1        -                None       -
      192.0.2.2    -                -
      No As-Path   -
u*>i Intra-Ad        192.0.2.3        100        0
      64496:1        -                None       -
      192.0.2.3    -                -
      No As-Path   -
u*>i Intra-Ad        192.0.2.4        100        0
      64496:1        -                None       -
      192.0.2.4    -                -
      No As-Path   -
-----
Routes : 4
=====
```

At this moment, there are no connected sources detected and no receivers wanting to join any multicast sources.

Each I-PMSI is seen as a PIM tunnel interface. As there are four routers in the MVPN, there are four I-PMSIs.

```
*A:PE-1# show router 1 pim tunnel-interface
=====
PIM Interfaces ipv4
=====
Interface      Originator Address  Adm  Opr  Transport Type
-----
mpls-if-73732  192.0.2.1          Up   Up   Tx-IPMSI
mpls-virt-if-1005857  192.0.2.3          Up   Up   Rx-IPMSI
mpls-virt-if-1005858  192.0.2.2          Up   Up   Rx-IPMSI
mpls-virt-if-1005859  192.0.2.4          Up   Up   Rx-IPMSI
-----
Interfaces : 4
=====
```

To be able to reach the source, a route for each source is included in the VRF for VPRN 1.

For PE-2, this looks as follows:

```
*A:PE-2# show router 1 route-table
=====
Route Table (Service: 1)
=====
Dest Prefix[Flags]                Type   Proto   Age           Pref
  Next Hop[Interface Name]         Metric
-----
10.10.10.1/32                     Remote BGP VPN 00h09m36s 170
      192.0.2.1 (tunneled)         2
10.10.10.3/32                     Remote BGP VPN 00h09m34s 170
      192.0.2.3 (tunneled)         2
172.16.1.0/24                     Remote BGP VPN 00h09m36s 170
      192.0.2.1 (tunneled)         2
172.16.2.0/24                     Local  Local   00h17m02s  0
      int-PE-2-H-2                  0
172.16.3.0/24                     Remote BGP VPN 00h09m34s 170
      192.0.2.3 (tunneled)         2
172.16.4.0/24                     Remote BGP VPN 00h08m21s 170
      192.0.2.4 (tunneled)         2
-----
No. of Routes: 6
Flags: n = Number of times nexthop is repeated
      B = BGP backup route available
      L = LFA nexthop available
      S = Sticky ECMP requested
=====
```

The sources at 172.16.1.0/24 and 172.16.3.0/24 are learned as BGP VPN routes.

The following output shows the BGP routes for these prefixes, for example for prefix 172.16.1.0/24 on PE-2:

```
*A:PE-2# show router bgp routes 172.16.1.0/24 vpn-ipv4 hunt
=====
BGP Router ID:192.0.2.2          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
Origin codes : i - IGP, e - EGP, ? - incomplete
=====
BGP VPN-IPv4 Routes
=====
RIB In Entries
-----
Network       : 172.16.1.0/24
Nexthop       : 192.0.2.1
Route Dist.   : 64496:1          VPN Label     : 524282
Path Id       : None
From          : 192.0.2.5
Res. Nexthop  : n/a
Local Pref.   : 100
Aggregator AS : None           Interface Name : int-PE-2-P-5
Atomic Aggr.  : Not Atomic     Aggregator    : None
AIGP Metric   : None           MED           : None
Connector     : None           IGP Cost      : 2
Community     : target:64496:1 l2-vpn/vrf-imp:192.0.2.1:2
               source-as:64496:0
=====
```



```

Cluster      : 1.1.1.1
Originator Id : 192.0.2.1      Peer Router Id : 192.0.2.5
Fwd Class    : None          Priority        : None
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 : 00h10m11s
VPRN Imported : 1
    
```

```

-----
RIB Out Entries
-----
    
```

```

-----
Routes : 1
=====
    
```

This prefix is advertised with three communities:

- A route target extended community
- An l2-vpn/vrf-import extended community.
- A source-AS extended community (not used in Intra-AS context).

The l2-vpn/vrf-import extended community is significant as it is a unique value. It represents a specific MVPN on a specific PE and comprises a 32 bit value that identifies the PE plus an index identifying the VRF. The 32 bit value is the system address. The index (2) can be derived from the command:

```

*A:PE-2# admin display-config index | match vprn1
      virtual-router "vprn1" 2 0
    
```

Therefore, the l2-vpn/vrf-import community for VPRN 1 on PE-1 is 192.0.2.1:2

This community attribute is included within the source-join BGP route that is sent in a BGP update by a receiver PE as it tries to join a multicast group with a source address that matches the 172.16.1.0/24 prefix. This ensures that the source-join route is only accepted as a valid route and imported by the PE that originated the source address prefix. This is described in the following section.

## Enable Redundant Sources

The redundant sources are now enabled so that multicast traffic flows into both PE-1 and PE-3, using groups (S-1,G-1) and (S-3,G-1), respectively.

On each of these PEs, a source active AD route is generated. By examining each receiver PE, these can be clearly seen.

For PE-2, the source active AD routes can be seen using the following command.

```

*A:PE-2# show router bgp routes mvpn-ipv4 type source-ad
=====
BGP Router ID:192.0.2.2      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
Origin codes : i - IGP, e - EGP, ? - incomplete
=====
BGP MVPN-IPv4 Routes
=====
Flag RouteType OriginatorIP LocalPref MED
RD SourceAS Path-Id IGP Cost
NextHop SourceIP Label
As-Path GroupIP
-----
u*>i Source-Ad - 200 0
64496:1 - None -
192.0.2.1 172.16.1.11
No As-Path 239.255.0.1
u*>i Source-Ad - 50 0
64496:1 - None -
192.0.2.3 172.16.3.33
No As-Path 239.255.0.1
-----
Routes : 2
=====

```

There are two routes present, one from each source for the same group from PE-1 and PE-3.

The PIM groups can now be seen on PE-1 as follows:

```

*A:PE-1# show router 1 pim group
=====
Legend: A = Active S = Standby
=====
PIM Groups ipv4
=====
Group Address Type Spt Bit Inc Intf No.0ifs
Source Address RP State Inc Intf(S)
-----
239.255.0.1 (S,G) int-PE-1-S-1 0
172.16.1.11 10.10.10.1
239.255.0.1 (S,G) 0
172.16.3.33 10.10.10.1
-----
Groups : 2
=====

```

There are two groups on PE-1. In addition to its locally connected source, PE-1 has also received a source active from PE-3 which has an incoming interface of the I-PMSI toward PE-3. The outgoing interface list is empty as there is no host wanting to become a group member.

Similarly, on the other sender, PE-3.

```

*A:PE-3# show router 1 pim group
=====
Legend: A = Active S = Standby
=====
PIM Groups ipv4
=====
Group Address Type Spt Bit Inc Intf No.0ifs
Source Address RP State Inc Intf(S)
-----
239.255.0.1 (S,G) 0
172.16.1.11 10.10.10.3
-----

```

```

239.255.0.1          (S,G)          int-PE-3-S-3    0
172.16.3.33         10.10.10.3
-----
Groups : 2
=====

```

By examining the receiver PE-2, it can be seen that the Source AD route for (S,G) (172.16.1.2, 239.255.0.1) from PE-1 has a higher local preference so it is chosen as the preferred (active) source. Examining these routes in more detail shows that each route is tagged with two communities: the route target extended community and the "red" or "blue" community, as seen in the following output.

```

*A:PE-2# show router bgp routes mvpn-ipv4 type source-ad hunt
=====
BGP Router ID:192.0.2.2      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP MVPN-IPv4 Routes
=====
RIB In Entries
-----
Route Type      : Source-Ad
Route Dist.     : 64496:1
Source IP       : 172.16.1.11
Group IP        : 239.255.0.1
Nexthop         : 192.0.2.1
Path Id         : None
From            : 192.0.2.5
Res. Nexthop    : 0.0.0.0
Local Pref.      : 200
Aggregator AS   : None
Atomic Aggr.    : Not Atomic
AIGP Metric     : None
Connector       : None
Community       : 64496:11 no-export target:64496:1
Cluster         : 1.1.1.1
Originator Id   : 192.0.2.1      Peer Router Id : 192.0.2.5
Origin          : IGP
Flags           : Used Valid Best
---snip---
VPRN Imported   : 1
-----

Route Type      : Source-Ad
Route Dist.     : 64496:1
Source IP       : 172.16.3.33
Group IP        : 239.255.0.1
Nexthop         : 192.0.2.3
Path Id         : None
From            : 192.0.2.5
Res. Nexthop    : 0.0.0.0
Local Pref.      : 50
Aggregator AS   : None
Atomic Aggr.    : Not Atomic
AIGP Metric     : None
Connector       : None
Community       : 64496:33 no-export target:64496:1
Interface Name  : NotAvailable
Aggregator      : None
MED             : 0
IGP Cost        : n/a

```

```

Cluster      : 1.1.1.1
Originator Id : 192.0.2.3          Peer Router Id : 192.0.2.5
Origin       : IGP
Flags        : Used Valid Best
---snip---
VPRN Imported : 1

-----
RIB Out Entries
-----
Routes : 2
=====

```

The local preference is set based on these community values.

A debug of the received BGP Source AD routes is as follows for PE-2:

```

2 2024/02/05 23:15:29.710 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.5
"Peer 1: 192.0.2.5: UPDATE
Peer 1: 192.0.2.5 - Received BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 90
  Flag: 0x90 Type: 14 Len: 29 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.1
    Type: Source-AD Len: 18 RD: 64496:1 Src: 172.16.1.11 Grp: 239.255.0.1
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 8 Community:
    64496:11
    no-export
  Flag: 0x80 Type: 9 Len: 4 Originator ID: 192.0.2.1
  Flag: 0x80 Type: 10 Len: 4 Cluster ID:
    1.1.1.1
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:64496:1
"

1 2024/02/05 23:15:29.418 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.5
"Peer 1: 192.0.2.5: UPDATE
Peer 1: 192.0.2.5 - Received BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 90
  Flag: 0x90 Type: 14 Len: 29 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.3
    Type: Source-AD Len: 18 RD: 64496:1 Src: 172.16.3.33 Grp: 239.255.0.1
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 8 Community:
    64496:33
    no-export
  Flag: 0x80 Type: 9 Len: 4 Originator ID: 192.0.2.3
  Flag: 0x80 Type: 10 Len: 4 Cluster ID:
    1.1.1.1
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:64496:1

```

"

Similarly, the Source Active routes on receiver PE-4 show that the highest local preference value of 200 is set for the SA route received from PE-3 with an originator ID of 192.0.2.3, as follows:

```
*A:PE-4# show router bgp routes mvpn-ipv4 type source-ad hunt
=====
BGP Router ID:192.0.2.4      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP MVPN-IPv4 Routes
=====
-----
RIB In Entries
-----
Route Type      : Source-Ad
Route Dist.     : 64496:1
Source IP       : 172.16.1.11
Group IP        : 239.255.0.1
Nextthop        : 192.0.2.1
Path Id         : None
From            : 192.0.2.5
Res. Nextthop   : 0.0.0.0
Local Pref.    : 50
Aggregator AS   : None
Atomic Aggr.    : Not Atomic
AIGP Metric     : None
Connector       : None
Community       : 64496:11 no-export target:64496:1
Cluster         : 1.1.1.1
Originator Id : 192.0.2.1
Origin          : IGP
Flags           : Used Valid Best
---snip---
VPRN Imported   : 1
-----

Route Type      : Source-Ad
Route Dist.     : 64496:1
Source IP       : 172.16.3.33
Group IP        : 239.255.0.1
Nextthop        : 192.0.2.3
Path Id         : None
From            : 192.0.2.5
Res. Nextthop   : 0.0.0.0
Local Pref.    : 200
Aggregator AS   : None
Atomic Aggr.    : Not Atomic
AIGP Metric     : None
Connector       : None
Community       : 64496:33 no-export target:64496:1
Cluster         : 1.1.1.1
Originator Id : 192.0.2.3
Origin          : IGP
Flags           : Used Valid Best
---snip---
VPRN Imported   : 1
-----
```

```
-----
RIB Out Entries
-----
Routes : 2
=====
```

## Host Group Membership

If the hosts then send a (\*,G) request to join the group, a source-join route is originated by each receiver PE toward the preferred source from the redundant list.

The following output shows a join originated by PE-2:

```
*A:PE-2# show debug
debug
  router "Base"
    bgp
      update
    exit
  exit
exit
```

```
3 2024/02/05 23:21:27.781 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.5
"Peer 1: 192.0.2.5: UPDATE
Peer 1: 192.0.2.5 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 84
  Flag: 0x90 Type: 14 Len: 33 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.2
    Type: Source-Join Len:22 RD: 64496:1 SrcAS: 64496 Src: 172.16.1.11 Grp: 239.255.0.1
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
  Flag: 0xc0 Type: 16 Len: 16 Extended Community:
    target:64496:1
    target:192.0.2.1:2
"
```

When an active source AD route is present, there is no shared join sent to the RP. Because the source address is known, only a source-join needs to be sent. The source-join is trying to become a member of group 239.255.0.1 with a source address of 172.16.1.11. As this is sent as a BGP routing update, this must be accepted by the MVPN VRF at the PE that originated the unicast route that represents the c-multicast source. As previously mentioned, there are two extended community values. The second of these is the I2-vpn/vrf-import route target for 192.0.2.1 (PE-1), so only PE-1 accepts this route.

Examining the PIM state table for PE-2 shows the presence of a group with multiple sources.

```
*A:PE-2# show router 1 pim group
```

```
=====
Legend:  A = Active   S = Standby
=====
```

```

PIM Groups ipv4
=====
Group Address          Type          Spt Bit  Inc Intf  No.0ifs
  Source Address      RP           State    Inc Intf(S)
-----
239.255.0.1           (*,G)                mpls-if-73729 1
*                      10.10.10.1
239.255.0.1           (S,G)          spt      mpls-if-73729 1
 172.16.1.11          10.10.10.1    A
239.255.0.1           (S,G)          S        mpls-if-73730 1
 172.16.3.33          10.10.10.1
-----
Groups : 3
=====

```

Each (S,G) has a state of either Active (A) or Standby (S), and the active group is chosen based on the Source Active AD with the highest local preference.

As a direct comparison, PE-4 also has the same two (S,G) states, but has a reversed active and standby source.

```

*A:PE-4# show router 1 pim group
=====
Legend:  A = Active  S = Standby
=====
PIM Groups ipv4
=====
Group Address          Type          Spt Bit  Inc Intf  No.0ifs
  Source Address      RP           State    Inc Intf(S)
-----
239.255.0.1           (*,G)                mpls-if-73730 1
*                      10.10.10.3
239.255.0.1           (S,G)          S        mpls-if-73729 1
 172.16.1.11          10.10.10.3
239.255.0.1           (S,G)          spt      mpls-if-73730 1
 172.16.3.33          10.10.10.3    A
-----
Groups : 3
=====

```

The Source Active ADs received on PE-4 have their local preference values based on the community string value.

```

*A:PE-4# show router bgp routes mvpn-ipv4 type source-ad hunt
=====
BGP Router ID:192.0.2.4      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP MVPN-IPv4 Routes
=====
RIB In Entries
-----
Route Type      : Source-Ad
Route Dist.     : 64496:1
Source IP       : 172.16.1.11

```

```

Group IP      : 239.255.0.1
NextHop      : 192.0.2.1
Path Id      : None
From         : 192.0.2.5
Res. NextHop : 0.0.0.0
Local Pref.  : 50
Aggregator AS : None
Atomic Aggr. : Not Atomic
AIGP Metric  : None
Connector    : None
Community    : 64496:11 no-export target:64496:1
Cluster      : 1.1.1.1
Originator Id : 192.0.2.1
Origin       : IGP
Flags        : Used Valid Best
---snip---
VPRN Imported : 1

```

```

-----
Route Type    : Source-Ad
Route Dist.   : 64496:1
Source IP     : 172.16.3.33
Group IP     : 239.255.0.1
NextHop      : 192.0.2.3
Path Id      : None
From         : 192.0.2.5
Res. NextHop : 0.0.0.0
Local Pref.  : 200
Aggregator AS : None
Atomic Aggr. : Not Atomic
AIGP Metric  : None
Connector    : None
Community    : 64496:33 no-export target:64496:1
Cluster      : 1.1.1.1
Originator Id : 192.0.2.3
Origin       : IGP
Flags        : Used Valid Best
Route Source  : Internal
---snip---
VPRN Imported : 1

```

-----  
RIB Out Entries  
-----

Routes : 2  
=====

## Sender PE MVPN Status

The MVPN status of the PE-1 sender PE is as follows:

```
*A:PE-1# show router 1 mvpn
```

```
=====
MVPN 1 configuration data
=====
```

```

signaling      : Bgp          auto-discovery   : Default
UMH Selection  : Highest-IP  SA withdrawn    : Disabled
intersite-shared : Enabled      Persist SA      : Enabled

```



```
vrf-import      : MVPN1_import
vrf-export      : MVPN1_export
vrf-target      : N/A
C-Mcast Import RT : target:192.0.2.1:2

ipmsi           : ldp
i-pmsi P2MP AdmSt : Up
i-pmsi Tunnel Name : mpls-if-73732
Mdt-type        : sender-only
ipmsi UMH RM    : Disabled

BSR signalling   : none
Wildcard s-pmsi : Disabled
Multistream-SPMSI : Disabled
s-pmsi           : none
data-delay-interval: 3 seconds
enable-asm-mdt   : N/A
spmsi UMH RM    : Disabled
```

The C-Mcast Import RT is set to <system-address>:<VPRN index>.

The VPRN index is derived from the following command:

```
*A:PE-1# admin display-config index | match vprn1
virtual-router "vprn1" 2 0
```

Any Source Join received must include this attribute along with the route target extended community. As previously stated, this is advertised within the VPN-IPv4 routes as a BGP attribute.

A source join received from PE-2 to join (S,G) (172.16.1.2, 239.255.0.1) is as follows:

```
*A:PE-1# show router bgp routes mvpn-ipv4 type source-join hunt
=====
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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP MVPN-IPv4 Routes
=====
-----
RIB In Entries
-----
Route Type      : Source-Join
Route Dist.     : 64496:1
Source AS       : 64496
Source IP      : 172.16.1.11
Group IP       : 239.255.0.1
Nexthop         : 192.0.2.2
Path Id         : None
From            : 192.0.2.5
Res. Nexthop    : 0.0.0.0
Local Pref.     : 100
Aggregator AS   : None
Aggregator      : None
Atomic Aggr.    : Not Atomic
AIGP Metric     : None
Connector       : None
Community     : no-export target:64496:1 target:192.0.2.1:2
Interface Name  : NotAvailable
MED             : 0
IGP Cost        : n/a
```

```

Cluster      : 1.1.1.1
Originator Id : 192.0.2.2      Peer Router Id : 192.0.2.5
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 : 00h08m11s
VPRN Imported : 1
    
```

-----  
RIB Out Entries  
-----

-----  
Routes : 1  
=====

The PIM status for this group on sender PE-1 is as follows:

```

*A:PE-1# show router 1 pim group 239.255.0.1 source 172.16.1.11 detail
=====
PIM Source Group ipv4
=====
Group Address      : 239.255.0.1
Source Address     : 172.16.1.11
RP Address         : 10.10.10.1
Advt Router       : 192.0.2.1
Flags              : spt                Type           : (S,G)
Mode               : sparse
MRIB Next Hop     : 172.16.1.11
MRIB Src Flags    : direct
Keepalive Timer Exp: 0d 00:02:00
Up Time           : 0d 00:14:28         Resolved By      : rtable-u

Up JP State       : Joined              Up JP Expiry     : 0d 00:00:00
Up JP Rpt        : Not Joined StarG     Up JP Rpt Override : 0d 00:00:00

Register State    : Pruned              Register Stop Exp : 0d 00:00:31
Reg From Anycast RP: No

Rpf Neighbor      : 172.16.1.11
Incoming Intf     : int-PE-1-S-1
Outgoing Intf List : mpls-if-73732

Curr Fwding Rate : 110.009 kbps
Forwarded Packets : 8683                Discarded Packets : 0
Forwarded Octets  : 11999906            RPF Mismatches    : 0
Spt threshold     : 0 kbps              ECMP opt threshold : 7
Admin bandwidth   : 1 kbps
    
```

-----  
Groups : 1  
=====

The outgoing interface list is the I-PMSI, and traffic is seen to be flowing because the current forwarding rate is non-zero.

Similarly for sender PE-3, the MVPN status is:

```
*A:PE-3# show router 1 mvpn
=====
MVPN 1 configuration data
=====
signaling          : Bgp          auto-discovery    : Default
UMH Selection      : Highest-IP   SA withdrawn     : Disabled
intersite-shared   : Enabled       Persist SA       : Enabled
vrf-import         : MVPN1_import
vrf-export         : MVPN1_export
vrf-target         : N/A
C-Mcast Import RT : target:192.0.2.3:2

ipmsi              : ldp
i-pmsi P2MP AdmSt : Up
i-pmsi Tunnel Name : mpls-if-73731
Mdt-type           : sender-only
ipmsi UMH RM       : Disabled

BSR signalling     : none
Wildcard s-pmsi   : Disabled
Multistream-SPMSI : Disabled
s-pmsi             : none
data-delay-interval: 3 seconds
enable-asm-mdt    : N/A
spmsi UMH RM       : Disabled
=====
```

The Source-Join route on PE-3 for this multicast group is:

```
*A:PE-3# show router bgp routes mvpn-ipv4 type source-join hunt
=====
BGP Router ID:192.0.2.3      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP MVPN-IPv4 Routes
=====
-----
RIB In Entries
-----
Route Type      : Source-Join
Route Dist.     : 64496:1
Source AS       : 64496
Source IP       : 172.16.3.33
Group IP        : 239.255.0.1
Nexthop         : 192.0.2.4
Path Id         : None
From            : 192.0.2.5
Res. Nexthop    : 0.0.0.0
Local Pref.     : 100
Aggregator AS   : None           Interface Name : NotAvailable
Atomic Aggr.    : Not Atomic     Aggregator     : None
AIGP Metric     : None           MED            : 0
Connector       : None           IGP Cost       : n/a
Community       : no-export target:64496:1 target:192.0.2.3:2
=====
```

```

Cluster      : 1.1.1.1
Originator Id : 192.0.2.4          Peer Router Id : 192.0.2.5
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 : 00h08m00s
VPRN Imported : 1
    
```

-----  
RIB Out Entries  
-----

-----  
Routes : 1  
=====

The PIM state for this group is as follows:

```

*A:PE-3# show router 1 pim group 239.255.0.1 source 172.16.3.33 detail
=====
PIM Source Group ipv4
=====
Group Address      : 239.255.0.1
Source Address     : 172.16.3.33
RP Address         : 10.10.10.3
Advt Router        : 192.0.2.3
Flags              : spt                Type           : (S,G)
Mode               : sparse
MRIB Next Hop     : 172.16.3.33
MRIB Src Flags    : direct
Keepalive Timer Exp: 0d 00:01:58
Up Time           : 0d 00:14:44          Resolved By      : rtable-u

Up JP State       : Joined                Up JP Expiry     : 0d 00:00:00
Up JP Rpt        : Not Joined StarG      Up JP Rpt Override : 0d 00:00:00

Register State    : Pruned                Register Stop Exp : 0d 00:00:43
Reg From Anycast RP: No

Rpf Neighbor      : 172.16.3.33
Incoming Intf     : int-PE-3-S-3
Outgoing Intf List : mpls-if-73731

Curr Fwding Rate  : 110.560 kbps
Forwarded Packets : 8838                  Discarded Packets : 0
Forwarded Octets  : 12214116              RPF Mismatches    : 0
Spt threshold     : 0 kbps                 ECMP opt threshold : 7
Admin bandwidth   : 1 kbps

-----
Groups : 1
=====
    
```

The preferred source remains active unless:

- The multicast source ceases to exist, the source PE withdraws the Source Active AD route
- Or a Source Active AD is received with a higher local preference.

## Conclusion

MVPN Source Redundancy provides an optimal solution for multicast routing in a VPRN. This protocol provides simple configuration, operation and guaranteed fast protection time. It could be used in a regionalized IPTV solution where multiple sources for the same TV channel are used.

## NG-MVPN Wildcard S-PMSI

This chapter provides information about next generation multicast virtual private networks (NG-MVPNs): use of wildcard selective PMSI.

Topics in this chapter include:

- [Applicability](#)
- [Overview](#)
- [Configuration](#)
- [Conclusion](#)

### Applicability

The chapter was initially written based on SR OS Release 13.0.R4, but the CLI in the current edition is based on SR OS Release 23.10.R1.

MPLS provider tunnels can be set up using multicast label distribution protocol (mLDP) or point-to-multipoint (P2MP) resource reservation protocol with traffic engineering (RSVP-TE) label switched paths (LSPs). SR OS Release 12.0.R4 or later is required for route reflectors (RRs) peering with multicast virtual private network (MVPN) PEs.

Provider multicast service interfaces (PMSIs) are signaled using mLDP. PE MVPN auto-discovery uses BGP MVPN IPv4 network layer routing.

Knowledge of multi-protocol BGP (MP-BGP), RFC 4364, *BGP/MPLS IP Virtual Private Networks (VPNs)*, RFC 6513, *Multicast in MPLS/BGP IP VPNs*/RFC 6514, *BGP Encodings and Procedures for Multicast in MPLS/BGP IP VPNs*, and RFC 6625, *Wildcards in Multicast VPN Auto-Discovery Routes*, is assumed throughout this chapter.

### Overview

Consider a service provider core network used to deliver multicast services to a number of receiver PEs using Next Generation MVPN techniques, as defined in RFC 6513/6514, where multicast traffic is forwarded between PEs across a mesh of provider tunnels.

The provider tunnel used is the default Inclusive PMSI (I-PMSI) that is instantiated between all source and receiver PEs. This results in a full mesh of provider tunnels between all PEs in the MVPN. In a large network, this can result in an inefficient use of bandwidth because multicast traffic is forwarded to all PEs regardless of whether the PE has an interested receiver. Some of the mesh scaling issues can be mitigated by using source-only/destination-only configuration on PEs. However, this technique requires additional configuration and is not fully optimal when mLDP is used in the core.

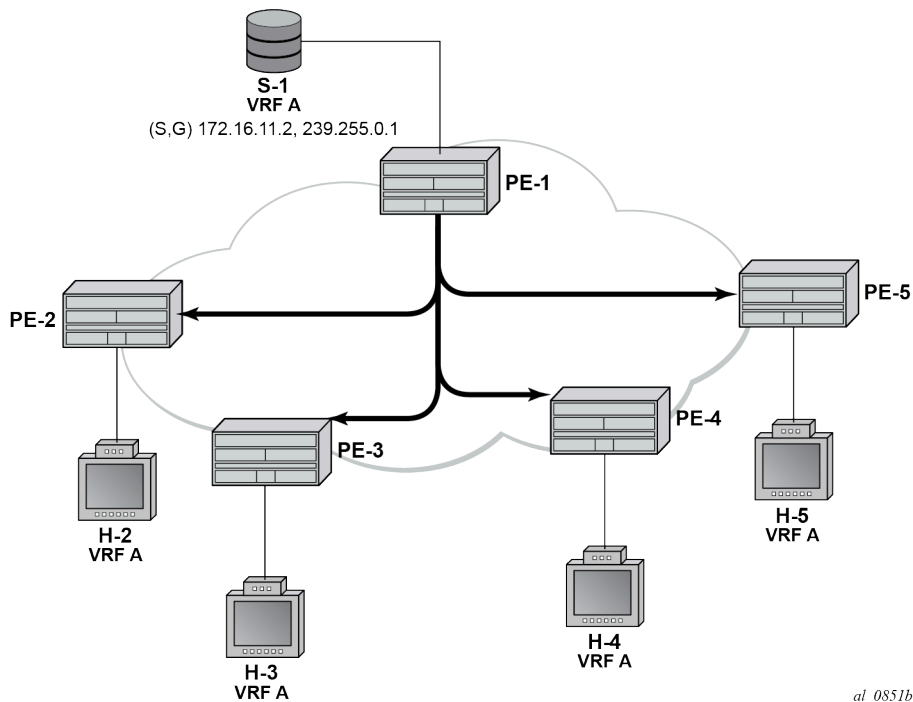
To address the preceding limitation, wildcard Selective PMSI (S-PMSI) has been developed, as per RFC 6625. In the standard customer signaling notation of (C-S,C-G), this becomes (C-\*,C-\*). Using methods defined in RFC 6625, it is possible to use a (C-\*,C-\*) S-PMSI as the default tunnel, where the receiver PE can join the S-PMSI by mapping the channel join to a wildcard channel group. Multiple channels can be

transported by the wildcard (C-\*,C-\*) S-PMSI, where an S-PMSI auto-discovery route is advertised with an empty channel group and source address:

1. Bandwidth savings can be achieved by the delivery of multicast channels on S-PMSIs, because traffic is not forwarded to PEs that have no interested receivers.
2. Control plane savings can be achieved by eliminating the need for the full tunnel mesh between all PEs. The wildcard S-PMSI is only signaled on PEs containing attached upstream multicast sources, for which the PE is resolved as an upstream multicast hop (UMH) within the MVPN.

Figure 61: Multicast VPN shows the concept of an MVPN.

Figure 61: Multicast VPN



In Figure 61: Multicast VPN, PE-1 has a directly connected multicast source (S-1). For clarity, consider this MVPN as a single multicast group. PE-1 is configured as a sender PE because it is the PE closest to the source. PE-2, PE-3, PE-4, and PE-5 are configured as receiver-only PEs because they have connected receiver hosts H-2, H-3, H-4, and H-5, respectively. Hosts H-2 to H-5 connected to receiver PEs can receive multicast channels from the source, S-1, connected to the source PE, PE-1, within the same virtual private routed network (VPRN).

Within the provider network, multicast traffic is delivered from the source PE to the receiver PE across a PMSI tunnel. This tunnel is, in this case, a P2MP LSP, with its root on PE-1 and with a leaf at each of the receiver PEs. Any traffic that is forwarded into the tunnel interface is replicated so that a single copy of a multicast stream is received at all PEs.

The PMSI tunnel is created after each PE has declared themselves as a member of the MVPN using BGP MVPN auto-discovery techniques.

There are two choices of PMSI:

- An I-PMSI, which is created on each PE within the MVPN, with a root at each PE and a leaf at all other PEs that are members of the MVPN. The I-PMSI is the default tunnel for all multicast traffic carried

between sender PE and receiver PEs. When at least one receiver PE has a host interested in becoming a member of a multicast group, traffic for that group is delivered to all PEs via the I-PMSI, regardless of whether they have an interested host. Receiver PEs with no such interested host then drop the traffic.

- An S-PMSI, which is created to carry multicast traffic to the subset of receiver PEs that have connected hosts interested in receiving multicast traffic. This can be for a specific group, so that one S-PMSI carries traffic for one multicast group, denoted as (C-S,C-G) or (C-\*,C-G). The S-PMSI can also be signaled to carry traffic for multiple multicast groups, denoted using a wildcard: (C-\*,C-\*) or (C-S,C-\*). The wildcard S-PMSI can be signaled in place of the I-PMSI, so that all traffic can be carried on the S-PMSI by default. In this case, no I-PMSI is signaled.

In the case of an I-PMSI, the tunnel type is included in the BGP auto-discovery intra-AD route originated and advertised to all other PEs within the VPRN.

If a wildcard S-PMSI is to be used and no I-PMSI tunnel is to be signaled, then an intra-AD route update for I-PMSI is advertised with no tunnel type attribute included. In addition, the source PE originates an additional S-PMSI auto-discovery route containing no source-encoding wildcard information.

[Table 4: S-PMSI Auto-Discovery BGP NLRI](#) shows the S-PMSI MVPN BGP network layer reachability information (NLRI) advertisement.

*Table 4: S-PMSI Auto-Discovery BGP NLRI*

Route Distinguisher (8 octets)
Multicast Source Length (1 octet)
Multicast Source (variable)
Multicast Group Length (1 octet)
Multicast Group (variable)
Originating Router IP Address

To signal the S-PMSI as wildcard (C-\*,C-\*) S-PMSI, the multicast source length and multicast group length fields are encoded with the value of zero (0), representing (C-\*,C-\*) wildcard.

The objectives are to:

- Configure multicast in a VPRN on PE-1 to PE-5 using mLDP as the tunnel signaling method.
- Connect multicast sources to the sender PE-1.
- Create receiver hosts that can receive multicast traffic from the source, and to observe the effect on the provider network.

The following configuration tasks should be completed as a prerequisite:

- Full mesh IS-IS or OSPF between each of the PE routers and the RR.
- Link-layer LDP between all PEs.
- mLDP used as the provider tunnel signaling protocol. This is enabled by default when link-layer LDP is enabled.

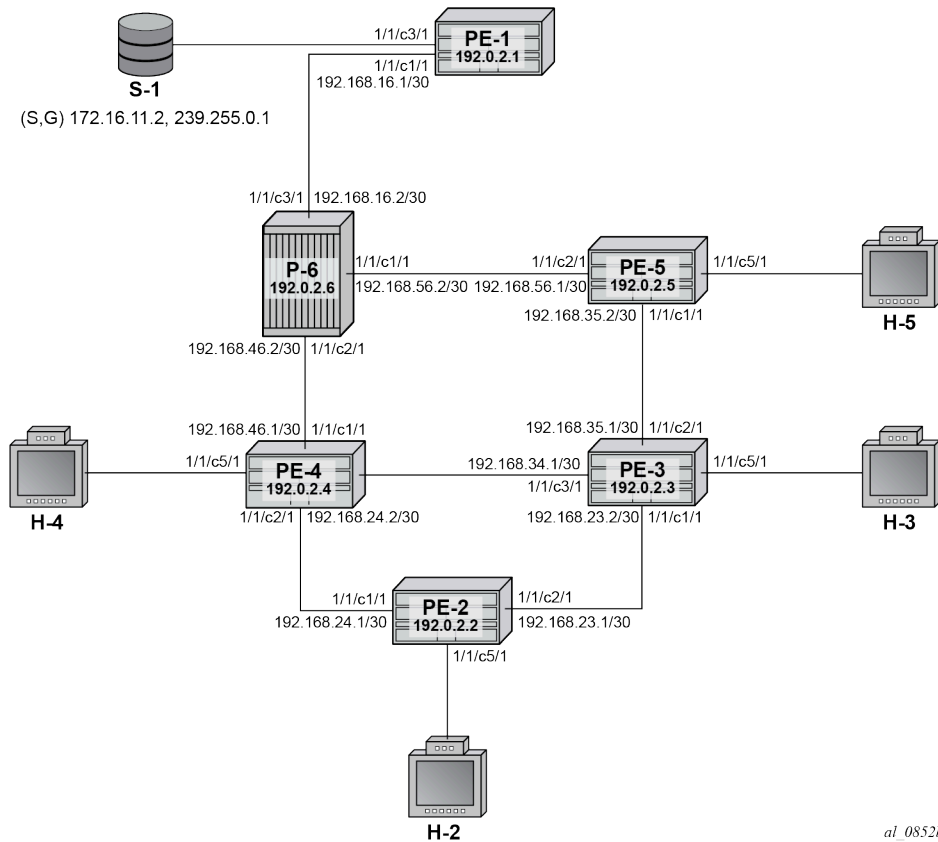
RSVP-TE is also supported as a provider tunnel signaling mechanism and could be used instead of mLDP.



## Configuration

The example topology is shown in [Figure 62: Schematic Topology](#), containing five PE routers. P-6 acts as an RR.

Figure 62: Schematic Topology



al\_0852b

## Global BGP Configuration

The first step is to configure an IBGP session between each of the PEs and the RR (PE-6) shown in [Figure 62: Schematic Topology](#). The address families negotiated between the IBGP peers are vpn-ipv4 (unicast routing) and mvpn-ipv4 (multicast routing).

The configuration for PE-1 is:

```
# on PE-1
configure
router
  autonomous-system 65545
  bgp
    group "INTERNAL"
      family vpn-ipv4 mvpn-ipv4
      type internal
      neighbor 192.0.2.6
```

```

        exit
    exit
exit all

```

The configuration for the other PE nodes is exactly the same.

The configuration for the RR on P-6 is:

```

# on P-6
configure
router
  autonomous-system 65545
  bgp
    cluster 0.0.0.1
    group "RR_CLIENTS"
      family vpn-ipv4 mvpn-ipv4
      type internal
      neighbor 192.0.2.1
      exit
      neighbor 192.0.2.2
      exit
      neighbor 192.0.2.3
      exit
      neighbor 192.0.2.4
      exit
      neighbor 192.0.2.5
      exit
    exit
  exit all

```

On PE-1, verify that the BGP session with RR on P-6 is established with address families vpn-ipv4 and mvpn-ipv4 capabilities negotiated:

```

*A:PE-1# show router bgp summary
=====
BGP Router ID:192.0.2.1      AS:65545      Local AS:65545
=====
BGP Admin State      : Up      BGP Oper State      : Up
Total Peer Groups    : 1        Total Peers          : 1
Total VPN Peer Groups : 0        Total VPN Peers      : 0
Current Internal Groups : 1      Max Internal Groups  : 1
Total BGP Paths       : 20       Total Path Memory    : 7360
---snip---
=====
BGP Summary
=====
Legend : D - Dynamic Neighbor
=====
Neighbor
Description
          AS PktRcvd InQ  Up/Down  State|Rcv/Act/Sent (Addr Family)
          PktSent OutQ
-----
192.0.2.6
          65545      11    0 00h03m43s 0/0/0 (VpnIPv4)
          11        0      0/0/0 (MvpnIPv4)
-----

```

The same command can be used on the other PEs to verify their BGP sessions to the RR.

## Configuring VPRN on PEs

The following outputs show the VPRN configurations for each PE. The specific MVPN configuration is shown later.

The VPRN configuration for PE-1 is:

```
# on PE-1
configure
  service
    vprn 1 name "VPRN 1" customer 1 create
    interface "int-PE-1-S-1" create
      address 172.16.11.1/24
      sap 1/1/c3/1 create
      exit
    exit
    interface "rp" create
      address 10.0.0.1/32
      loopback
    exit
    bgp-ipvpn
      mpls
        auto-bind-tunnel
        resolution-filter
        ldp
        exit
        resolution filter
      exit
      route-distinguisher 65545:1
      vrf-target target:65545:1
      no shutdown
    exit
  exit
  pim
    apply-to all
    rp
      static
        address 10.0.0.1
        group-prefix 239.255.0.0/16
      exit
    exit
  exit
  no shutdown
exit
no shutdown
exit all
```

There is a single interface toward S-1 from which the multicast group is transmitted.

If the customer signaling uses PIM ASM, a customer Rendezvous Point (RP) is required.

A loopback interface called "rp" acts as the RP for all group prefixes in the 239.255.0.0/16 range. This is the RP for all groups.

MVPN configuration enables BGP as both the auto-discovery mechanism and the customer multicast signaling protocol across the VPRN. The provider tunnel between PEs within the MVPN is signaled using mLDP.

PE-2 contains an attached receiver, therefore a single interface is configured to accommodate this, as follows. The RP is configured as a static RP:

```
# on PE-2
configure
service
  vprn 1 name "VPRN 1" customer 1 create
  interface "int-PE-2-H-2" create
    address 172.16.22.1/24
    sap 1/1/c5/1 create
  exit
exit
bgp-ipvpn
  mpls
    auto-bind-tunnel
    resolution-filter
    ldp
  exit
  resolution filter
  exit
  route-distinguisher 65545:1
  vrf-target target:65545:1
  no shutdown
exit
exit
igmp
  interface "int-PE-2-H-2"
    no shutdown
  exit
  no shutdown
exit
pim
  apply-to all
  rp
    static
    address 10.0.0.1
    group-prefix 239.255.0.0/16
  exit
  exit
  no shutdown
exit
  no shutdown
exit all
```

PE-3 also has an attached receiver, as follows:

```
# on PE-3
configure
service
  vprn 1 name "VPRN 1" customer 1 create
  interface "int-PE-3-H-3" create
    address 172.16.33.1/24
    sap 1/1/c5/1 create
  exit
exit
bgp-ipvpn
  mpls
    auto-bind-tunnel
    resolution-filter
    ldp
  exit
  resolution filter
```

```

        exit
        route-distinguisher 65545:1
        vrf-target target:65545:1
        no shutdown
    exit
exit
igmp
    interface "int-PE-3-H-3"
        no shutdown
    exit
    no shutdown
exit
pim
    apply-to all
    rp
        static
            address 10.0.0.1
            group-prefix 239.255.0.0/16
        exit
    exit
    no shutdown
exit
    no shutdown
exit all

```

The configuration for PE-4 and PE-5 is similar.

## MVPN Configuration for PEs

The provider-tunnel inclusive configuration specifies that a wildcard S-PMSI is used instead of an I-PMSI as the default PMSI. The MVPN configuration for all PEs is:

```

# on all PEs
configure
    service
        vprn 1
            mvpn
                auto-discovery default
                c-mcast-signaling bgp
                provider-tunnel
                    inclusive
                    mldp
                        no shutdown
                    exit
                wildcard-spmsi
            exit
        exit
        vrf-target unicast
    exit
    no shutdown
exit all

```

The keyword **wildcard-spmsi** reduces the number of PMSIs signaled. If there are no group sources on the receiver PEs, there are no S-PMSI signaled. This has an effect similar to configuring the receiver PEs as MDT-type receiver-only.

## Provider Tunnel Signaling

Each PE originates BGP MVPN intra-AD routes to determine membership of an MVPN.

The provider tunnels constructed between the PEs within the VPRN are signaled on receipt of an intra-AD route update from other PEs. The intra-AD update message contains details of the originator, along with the VRF route-target extended community. If an I-PMSI is to be signaled, a PMSI tunnel attribute is included that determines the tunnel type, such as LDP P2MP LSP. PEs that receive this intra-AD update import the route into the MVPN, then signal a P2MP LDP label map toward the originator, which is the root of the LDP P2MP LSP.

However, if a wildcard S-PMSI is to be used as the default PMSI, no PMSI tunnel attribute is included within the intra-AD update.

The following output shows a BGP update originated by PE-1, and received by PE-2:

```
*A:PE-2# show router bgp routes mvpn-ipv4 type intra-ad rd 65545:1 detail originator-ip
192.0.2.1
=====
BGP Router ID:192.0.2.2      AS:65545      Local AS:65545
=====
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 MVPN-IPv4 Routes
=====
Original Attributes

Route Type      : Intra-Ad
Route Dist.     : 65545:1
Originator IP   : 192.0.2.1
NextHop         : 192.0.2.1
Path Id         : None
From            : 192.0.2.6
Res. NextHop    : 0.0.0.0
Local Pref.     : 100
Aggregator AS   : None
Atomic Aggr.    : Not Atomic
AIGP Metric     : None
Connector       : None
Community       : no-export target:65545:1
Cluster         : 0.0.0.1
Originator Id   : 192.0.2.1
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
Source Class    : 0
Add Paths Send  : Default
Last Modified   : 00h01m19s
VPRN Imported   : 1

Interface Name  : NotAvailable
Aggregator     : None
MED             : 0
IGP Cost        : n/a

Peer Router Id  : 192.0.2.6
Final Orig Val  : N/A
Dest Class      : 0

Modified Attributes

Route Type      : Intra-Ad
```

```

Route Dist.      : 65545:1
---snip---
VPRN Imported   : 1

-----
Routes : 1
=====

```

There is no PMSI tunnel attribute included, and the route is imported into the correct VPRN (VPRN 1).

The intra-AD originated by PE-2 is:

```

*A:PE-1# show router bgp routes mvpn-ipv4 type intra-ad rd 65545:1 originator-ip 192.0.2.2
hunt
=====
BGP Router ID:192.0.2.1      AS:65545      Local AS:65545
=====
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 MVPN-IPv4 Routes
=====
RIB In Entries
-----
Route Type      : Intra-Ad
Route Dist.     : 65545:1
Originator IP   : 192.0.2.2
Nextthop       : 192.0.2.2
Path Id        : None
From           : 192.0.2.6
Res. Nextthop  : 0.0.0.0
Local Pref.    : 100
Aggregator AS  : None
Atomic Aggr.   : Not Atomic
AIGP Metric    : None
Connector      : None
Community      : no-export target:65545:1
Cluster        : 0.0.0.1
Originator Id  : 192.0.2.2
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
Source Class   : 0
Add Paths Send : Default
Last Modified  : 00h01m43s
VPRN Imported  : 1

-----
RIB Out Entries
-----
Routes : 1
=====

```

This output also contains no PMSI tunnel attribute: PE-2 has no group source and there is no S-PMSI signaled. All other receiver PEs originate a similar intra-AD route.

The following output shows all intra-AD routes originated by the PEs within the VPRN, as received by PE-1:

```
*A:PE-1# show router bgp routes mvpn-ipv4 type intra-ad rd 65545:1
=====
BGP Router ID:192.0.2.1      AS:65545      Local AS:65545
=====
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 MVPN-IPv4 Routes
=====
Flag  RouteType      OriginatorIP      LocalPref  MED
      RD           SourceAS          Path-Id     IGP Cost
      Nexthop      SourceIP          Label
      As-Path      GroupIP
-----
i     Intra-Ad        192.0.2.1        100        0
      65545:1        -                None        -
      192.0.2.1      -                -
      No As-Path    -
u*>i Intra-Ad        192.0.2.2        100        0
      65545:1        -                None        -
      192.0.2.2      -                -
      No As-Path    -
u*>i Intra-Ad        192.0.2.3        100        0
      65545:1        -                None        -
      192.0.2.3      -                -
      No As-Path    -
u*>i Intra-Ad        192.0.2.4        100        0
      65545:1        -                None        -
      192.0.2.4      -                -
      No As-Path    -
u*>i Intra-Ad        192.0.2.5        100        0
      65545:1        -                None        -
      192.0.2.5      -                -
      No As-Path    -
-----
Routes : 5
=====
```

Instead of an I-PMSI being signaled, an S-PMSI AD route update is advertised by PE-1 to all receiver PEs within the MVPN. The NLRI encoding has a zero length field for both source and group addresses, so is seen to represent multicast group (C-\*,C-\*). This is wildcard nomenclature for both source and group addresses.

The BGP route as advertised by PE-1:

```
*A:PE-1# show router bgp routes mvpn-ipv4 type spmsi-ad rd 65545:1 hunt
=====
BGP Router ID:192.0.2.1      AS:65545      Local AS:65545
=====
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 MVPN-IPv4 Routes
=====
-----
RIB In Entries
-----
---snip---
-----
RIB Out Entries
-----
Route Type      : Spmsi-Ad
Route Dist.     : 65545:1
Originator IP   : 192.0.2.1
Source IP       : 0.0.0.0
Group IP        : 0.0.0.0
Nextthop       : 192.0.2.1
Path Id        : None
To             : 192.0.2.6
Res. Nextthop  : n/a
Local Pref.    : 100
Aggregator AS  : None
Interface Name : NotAvailable
Aggregator     : None
---snip---
Community      : no-export target:65545:1
Cluster       : No Cluster Members
Originator Id  : None
Peer Router Id : 192.0.2.6
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
-----
PMSI Tunnel Attributes :
Tunnel-type   : LDP P2MP LSP
Flags        : Type: RNVE(0) BM: 0 U: 0 Leaf: not required
MPLS Label   : 0
Root-Node    : 192.0.2.1
LSP-ID       : 8193
-----
-----
Routes : 6
=====

```

The source IP and group IP address fields are advertised as 0.0.0.0, and the tunnel type attribute is now included as an LDP P2MP LSP.

The following output shows the MVPN status on PE-1, with the I-PMSI tunnel name containing a wildcard S-PMSI denoted by (W):

```

*A:PE-1# show router 1 mvpn
=====
MVPN 1 configuration data
=====
signaling      : Bgp
UMH Selection  : Highest-Ip
auto-discovery : Default
intersite-shared : Enabled
SA withdrawn   : Disabled
Persist SA    : Disabled
vrf-import    : N/A
vrf-export    : N/A
vrf-target    : unicast
C-Mcast Import RT : target:192.0.2.1:2

```

```

ipmsi                               : ldp
i-pmsi P2MP AdmSt                   : Up
i-pmsi Tunnel Name                 : mpls-if-73728(W)
Mdt-type                             : sender-receiver
ipmsi UMH RM                         : Disabled

BSR signalling                       : none
Wildcard s-pmsi                   : Enabled
Multistream-SPMSI                   : Disabled
s-pmsi                             : none
data-delay-interval                 : 3 seconds
enable-asm-mdt                      : N/A
spmsi UMH RM                        : Disabled
=====

```

At this point, there is no interested host and no customer multicast flow (c-flow), so there is no S-PMSI LDP P2MP LSP signaled.

## Data Transmission at Source PE

Multicast traffic for a particular group is forwarded between the sender and receiver PE over a provider tunnel (PMSI). Because there is no default I-PMSI present, the receiver PE has to choose an S-PMSI to be used for forwarding, based on the NLRI contained within the S-PMSI AD routes.

The provider tunnel is signaled using a P2MP LDP label mapping message toward the root address signaled in the wildcard S-PMSI AD BGP update message. As previously shown, the update message is based on whether traffic is being forwarded on the shared or source-based shortest path tree.

When joining the shared tree, a c-multicast shared-join is sent toward the appropriate PE, which represents the UMH toward the RP. The UMH is chosen from the unicast route that represents the RP address.

When joining the shortest path tree, a source-join c-multicast route is sent toward the UMH chosen from the unicast route that represents the actual source address. In both cases, the VPN-IPv4 unicast route advertises a VRF route import community that contains the system address as a next hop. Upon receipt of these updates, the UMH PE forwards traffic along the signaled wildcard S-PMSI.

Each S-PMSI is bound to one or more flows, as determined by the NLRI contained within the S-PMSI BGP update. The use of the wildcard within the BGP update to replace the c-source and c-group allows multiple flows to be bound to a single provider tunnel.

Traffic is only forwarded upon reception of a BGP MVPN source-join or shared-join BGP route at the sender PE.

## Data Reception at Receiver PE

When the sender PE has originated an S-PMSI AD route update, each receiver PE installs the route into its VRF. The S-PMSIs installed are used to select an appropriate upstream multicast router for a c-flow when an attached receiver is interested in receiving traffic from that c-flow.

The receiver PE receives a flow based on the best match of the incoming (C-S,C-G) or (C-\*,C-G) IGMP/MLD or PIM join.

If an IGMP/MLD group membership query or PIM join is received by the receiver PE over an attachment circuit for a group, the contained (C-S,C-G) or (C-\*,C-G) must match the (C-S,C-G) contained within any installed S-PMSI AD route. In the case of the wildcard S-PMSI being the only installed NLRI, this is a match; that is, incoming (C-\*,C-G) or (C-S,C-G) match the S-PMSI (C-\*,C-\*) .

In this example, the c-group flow is 239.255.0.1.

## Traffic Flow

A traffic stream representing a c-flow is enabled on PE-1: group address 239.255.0.1 with source address of 172.16.11.2. The RP for this group is found locally on PE-1. The outgoing interface list is empty:

```
*A:PE-1# show router 1 pim group detail
=====
PIM Source Group ipv4
=====
Group Address      : 239.255.0.1
Source Address     : 172.16.11.2
RP Address         : 10.0.0.1
Advt Router       : 192.0.2.1
Flags              :
Mode               : sparse
MRIB Next Hop     : 172.16.11.2
MRIB Src Flags    : direct
Keepalive Timer Exp: 0d 00:02:47
Up Time           : 0d 00:00:43      Resolved By      : rtable-u

Up JP State        : Not Joined      Up JP Expiry     : 0d 00:00:00
Up JP Rpt         : Not Joined StarG Up JP Rpt Override: 0d 00:00:00

Register State    : Pruned           Register Stop Exp: 0d 00:01:14
Reg From Anycast RP: No

Rpf Neighbor      : 172.16.11.2
Incoming Intf     : int-PE-1-S-1
Outgoing Intf List:

Curr Fwding Rate  : 9627.528 kbps
Forwarded Packets : 52083             Discarded Packets : 0
Forwarded Octets  : 51145506        RPF Mismatches    : 0
Spt threshold     : 0 kbps           ECMP opt threshold: 7
Admin bandwidth   : 1 kbps
-----
Groups : 1
=====
```

A host connected to PE-2 sends a (C-\*,C-G) IGMPv2 group membership query for group 239.255.0.1.

PE-2 sends a BGP MVPN shared-join route update toward PE-1, where the RP address of the group 10.0.0.1 is found.

The following debug output from PE-2 shows the shared-join BGP route update transmitted by PE-2:

```
1 2023/12/07 12:18:46.927 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.6
"Peer 1: 192.0.2.6: UPDATE
Peer 1: 192.0.2.6 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 76
  Flag: 0x90 Type: 14 Len: 33 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.2
    Type: Shared-Join Len:22 RD: 65545:1 SrcAS: 65545 Src: 10.0.0.1 Grp: 239.255.0.1
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
```

```

Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
Flag: 0xc0 Type: 8 Len: 4 Community:
      no-export
Flag: 0xc0 Type: 16 Len: 8 Extended Community:
      target:192.0.2.1:2
"

```

Upon receipt of the shared-join, traffic flows on the shared tree toward the receiver PE. This flows on the default wildcard S-PMSI, as shown in the outgoing interface list:

```

*A:PE-1# show router 1 pim group 239.255.0.1 type starg detail
=====
PIM Source Group ipv4
=====
Group Address      : 239.255.0.1
Source Address     : *
RP Address         : 10.0.0.1
Advt Router        : 192.0.2.1
Flags              :                               Type           : (*,G)
Mode               : sparse
MRIB Next Hop     :
MRIB Src Flags     : self
Keepalive Timer    : Not Running
Up Time            : 0d 00:05:34      Resolved By         : rtable-u

Up JP State        : Joined              Up JP Expiry        : 0d 00:00:26
Up JP Rpt          : Not Joined StarG    Up JP Rpt Override  : 0d 00:00:00

Rpf Neighbor       :
Incoming Intf      :
Outgoing Intf List : mpls-if-73728(W)

Curr Fwding Rate   : 0.000 kbps
Forwarded Packets  : 0                  Discarded Packets   : 0
Forwarded Octets   : 0                  RPF Mismatches      : 0
Spt threshold      : 0 kbps              ECMP opt threshold  : 7
Admin bandwidth    : 1 kbps
-----
Groups : 1
=====

```

When traffic is received on the shared tree by PE-2, the source address is learned, so a source-join BGP route update is sent toward the UMH PE, which contains the source address of 172.16.11.2. The UMH is chosen from the unicast route-table using a lookup for the best route matching the source address.

The following debug output from PE-2 shows the BGP source-join route update toward the source of group 239.255.0.1, as transmitted by PE-2:

```

3 2023/12/07 12:18:46.938 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.6
"Peer 1: 192.0.2.6: UPDATE
Peer 1: 192.0.2.6 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 76
  Flag: 0x90 Type: 14 Len: 33 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.2
    Type: Source-Join Len:22 RD: 65545:1 SrcAS: 65545 Src: 172.16.11.2 Grp: 239.255.0.1
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100

```

```

Flag: 0xc0 Type: 8 Len: 4 Community:
no-export
Flag: 0xc0 Type: 16 Len: 8 Extended Community:
target:192.0.2.1:2
"

```

The c-flow toward host H-2 flows along the shortest path tree, and on PE-1 the outgoing interface list is populated with the wildcard S-PMSI:

```

*A:PE-1# show router 1 pim group detail 239.255.0.1 source 172.16.11.2
=====
PIM Source Group ipv4
=====
Group Address      : 239.255.0.1
Source Address     : 172.16.11.2
RP Address         : 10.0.0.1
Advt Router        : 192.0.2.1
Flags              : spt, rpt-prn-des   Type           : (S,G)
Mode               : sparse
MRIB Next Hop     : 172.16.11.2
MRIB Src Flags    : direct
Keepalive Timer Exp: 0d 00:01:05
Up Time           : 0d 00:14:28          Resolved By      : rtable-u

Up JP State        : Joined              Up JP Expiry     : 0d 00:00:00
Up JP Rpt         : Pruned              Up JP Rpt Override: 0d 00:00:00

Register State    : Pruned              Register Stop Exp : 0d 00:00:22
Reg From Anycast RP: No

Rpf Neighbor      : 172.16.11.2
Incoming Intf     : int-PE-1-S-1
Outgoing Intf List : mpls-if-73728(W)

Curr Fwding Rate  : 9368.280 kbps
Forwarded Packets : 1050618              Discarded Packets : 0
Forwarded Octets  : 1031706876          RPF Mismatches    : 0
Spt threshold     : 0 kbps              ECMP opt threshold: 7
Admin bandwidth   : 1 kbps
-----
Groups : 1
=====

```

The outgoing interface is the MPLS interface mpls-if-73728. This maps to a P2MP LDP label binding from which the p2mp-id can be derived:

```

*A:PE-1# show router ldp bindings active p2mp ipv4 opaque-type generic
=====
LDP Bindings (IPv4 LSR ID 192.0.2.1)
(IPv6 LSR ID ::)
=====
Label Status:
  U - Label In Use, N - Label Not In Use, W - Label Withdrawn
  WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
  e - Label ELC
FEC Flags:
  LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
  BA - ASBR Backup FEC
=====
LDP Generic IPv4 P2MP Bindings (Active)

```

```

=====
P2MP-Id                               Interface
RootAddr                               Op
IngLbl                                 EgrLbl
EgrNH                                  EgrIf/LspId
-----
8193                                  73728
192.0.2.1                             Push
--                                       524281
192.168.16.2                           1/1/c1/1
-----
No. of Generic IPv4 P2MP Active Bindings: 1
=====

```

After the source-join is received, the sender PE advertises a source-active AD BGP route to all PEs within the MVPN, to announce the presence of a (C-S,C-G) group. The following debug output shows the source-active AD route received on PE-2:

```

5 2023/12/07 12:18:46.946 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.6
"Peer 1: 192.0.2.6: UPDATE
Peer 1: 192.0.2.6 - Received BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 86
  Flag: 0x90 Type: 14 Len: 29 Multiprotocol Reachable NLRI:
    Address Family MVPN_IPV4
    NextHop len 4 NextHop 192.0.2.1
    Type: Source-AD Len: 18 RD: 65545:1 Src: 172.16.11.2 Grp: 239.255.0.1
  Flag: 0x40 Type: 1 Len: 1 Origin: 0
  Flag: 0x40 Type: 2 Len: 0 AS Path:
  Flag: 0x80 Type: 4 Len: 4 MED: 0
  Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
  Flag: 0xc0 Type: 8 Len: 4 Community:
    no-export
  Flag: 0x80 Type: 9 Len: 4 Originator ID: 192.0.2.1
  Flag: 0x80 Type: 10 Len: 4 Cluster ID:
    0.0.0.1
  Flag: 0xc0 Type: 16 Len: 8 Extended Community:
    target:65545:1
"

```

The PIM status of the group on receiver PE-2 shows that the incoming interface is the wildcard S-PMSI originated on PE-1, as denoted by the (W):

```

*A:PE-2# show router 1 pim group 239.255.0.1 source 172.16.11.2 detail
=====
PIM Source Group ipv4
=====
Group Address      : 239.255.0.1
Source Address     : 172.16.11.2
RP Address         : 10.0.0.1
Advt Router       : 192.0.2.1
Flags              : spt                               Type           : (S,G)
Mode               : sparse
MRIB Next Hop     : 192.0.2.1
MRIB Src Flags    : remote
Keepalive Timer Exp: 0d 00:00:24
Up Time           : 0d 00:06:37                       Resolved By     : rtable-u

Up JP State       : Joined                             Up JP Expiry    : 0d 00:00:22
Up JP Rpt        : Not Pruned                         Up JP Rpt Override : 0d 00:00:00

```

```

Register State      : No Info
Reg From Anycast RP: No

Rpf Neighbor       : 192.0.2.1
Incoming Intf    : mpls-if-73729(W)
Outgoing Intf List : int-PE-2-H-2

Curr Fwding Rate   : 9332.928 kbps
Forwarded Packets  : 472187          Discarded Packets : 0
Forwarded Octets   : 463687634       RPF Mismatches    : 0
Spt threshold      : 0 kbps           ECMP opt threshold: 7
Admin bandwidth    : 1 kbps

-----
Groups : 1
=====

```

The S-PMSI is an LDP P2MP LSP. The LDP label binding for P2MP LSP-Id 8193 on PE-2 shows that the label operation is a label pop:

```

*A:PE-2# show router ldp bindings active p2mp p2mp-id 8193 root 192.0.2.1

=====
LDP Bindings (IPv4 LSR ID 192.0.2.2)
(IPv6 LSR ID ::)
=====
Label Status:
  U - Label In Use, N - Label Not In Use, W - Label Withdrawn
  WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
  e - Label ELC
FEC Flags:
  LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
  BA - ASBR Backup FEC
=====
LDP Generic IPv4 P2MP Bindings (Active)
=====
P2MP-Id          Interface
RootAddr         Op
IngLbl           EgrLbl
EgrNH            EgrIf/LspId
-----
8193            73729
192.0.2.1      Pop
524280         --
  --             --
-----
No. of Generic IPv4 P2MP Active Bindings: 1
=====

```

PE-3 has no host joined to c-flow group 239.255.0.1. However, it contains the PIM state for this group because of the presence of the source-active AD route within the VRF. This route was received when the host connected to PE-2 joined the c-flow group.

The following output shows the source-active AD route within PE-3 for group 239.255.0.1 from source 172.16.11.2:

```

*A:PE-3# show router bgp routes mvpn-ipv4 type source-ad rd 65545:1

=====
BGP Router ID:192.0.2.3      AS:65545      Local AS:65545
=====
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 MVPN-IPv4 Routes

=====

Flag	RouteType	OriginatorIP	LocalPref	MED
	RD	SourceAS	Path-Id	IGP Cost
	Nexthop	SourceIP		Label
	As-Path	GroupIP		
-----				
u*>i	Source-Ad	-	100	0
	65545:1	-	None	-
	192.0.2.1	172.16.11.2		
	No As-Path	239.255.0.1		
-----				

Routes : 1

=====

However, traffic is not received from the S-PMSI because there is no label binding for the LDP P2MP LSP. The following output shows that there is no label binding for the LSP Id 8193, which has its root on PE-1:

```
*A:PE-3# show router ldp bindings p2mp p2mp-id 8193 root 192.0.2.1
```

```
=====
LDP Bindings (IPv4 LSR ID 192.0.2.3)
(IPv6 LSR ID ::)
```

Label Status:

```
U - Label In Use, N - Label Not In Use, W - Label Withdrawn
WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
e - Label ELC
```

FEC Flags:

```
LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
BA - ASBR Backup FEC
```

=====

LDP Generic IPv4 P2MP Bindings

=====

P2MP-Id	Interface
RootAddr	
Peer	
IngLbl	EgrLbl
EgrNH	EgrIf/LspId

No Matching Entries Found

=====

A static IGMPv2 (\*,G) group is created on interface int-PE-3-H3 for group 239.255.0.1 toward PE-3. The following debug output shows the process.

```
5 2023/12/07 12:27:42.918 UTC MINOR: DEBUG #2001 vprn1 IGMP[vprn1 inst 2]
"IGMP[vprn1 inst 2]: igmpIfSGStaticAdd
Adding Static SG (0.0.0.0,239.255.0.1) to IGMP interface int-PE-3-H-3 [ifIndex 5]"

6 2023/12/07 12:27:42.918 UTC MINOR: DEBUG #2001 vprn1 IGMP[vprn1 inst 2]
"IGMP[vprn1 inst 2]: igmpIfGroupAdd
Adding 239.255.0.1 to IGMP interface int-PE-3-H-3 [ifIndex 5] database"

7 2023/12/07 12:27:42.918 UTC MINOR: DEBUG #2001 vprn1 IGMP[vprn1 inst 2]
"IGMP[vprn1 inst 2]: igmpProcessGroupRec
Process group rec CHG_TO_EXCL received on interface int-PE-3-H-3 [ifIndex 5]"
```



```
for group 239.255.0.1 in mode INCLUDE. Num srcs 0"
```

```
8 2023/12/07 12:27:42.918 UTC MINOR: DEBUG #2001 vprn1 IGMP[vprn1 inst 2]
"IGMP[vprn1 inst 2]: igmpIfSrcAdd
Adding i/f source entry for interface int-PE-3-H-3 [ifIndex 5] (*,239.255.0.1) to
IGMP fwdList Database, redir if N/A"
```

A similar process takes place when receiver host H-3 sends an unsolicited IGMPv2 group membership query for this group. The first message would correspond to the IGMP query instead.

After the IGMP interface source entry has been added for interface int-PE-3-H-3, an mLDP P2MP label mapping message is sent from PE-3 toward the root node PE-1, as follows:

```
9 2023/12/07 12:27:42.918 UTC MINOR: DEBUG #2001 Base LDP
"LDP: Binding
Sending Label mapping label 524280 for P2MP: root = 192.0.2.1, T: 1, L: 4,
TunnelId: 8193 to peer 192.0.2.4:0."

10 2023/12/07 12:27:42.918 UTC MINOR: DEBUG #2001 Base LDP
"LDP: LDP
Send Label Mapping packet (msgId 239) to 192.0.2.4:0
Protocol version = 1
Label 524280 advertised for the following FECs
P2MP: root = 192.0.2.1, T: 1, L: 4, TunnelId: 8193
"
```

BGP shared-join and source-join BGP route updates are sent via the RR toward the RP (source = 10.0.0.1) and the actual source (172.16.11.2), respectively:

```
11 2023/12/07 12:27:42.918 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.6
"Peer 1: 192.0.2.6: UPDATE
Peer 1: 192.0.2.6 - Send BGP UPDATE:
Withdrawn Length = 0
Total Path Attr Length = 100
Flag: 0x90 Type: 14 Len: 57 Multiprotocol Reachable NLRI:
Address Family MVPN_IPV4
NextHop len 4 NextHop 192.0.2.3
Type: Source-Join Len:22 RD: 65545:1 SrcAS: 65545 Src: 172.16.11.2 Grp: 239.255.0.1
Type: Shared-Join Len:22 RD: 65545:1 SrcAS: 65545 Src: 10.0.0.1 Grp: 239.255.0.1
Flag: 0x40 Type: 1 Len: 1 Origin: 0
Flag: 0x40 Type: 2 Len: 0 AS Path:
Flag: 0x80 Type: 4 Len: 4 MED: 0
Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
Flag: 0xc0 Type: 8 Len: 4 Community:
no-export
Flag: 0xc0 Type: 16 Len: 8 Extended Community:
target:192.0.2.1:2
"
```

The PIM status for the c-group 239.255.0.1 on PE-3 is as follows:

```
*A:PE-3# show router 1 pim group

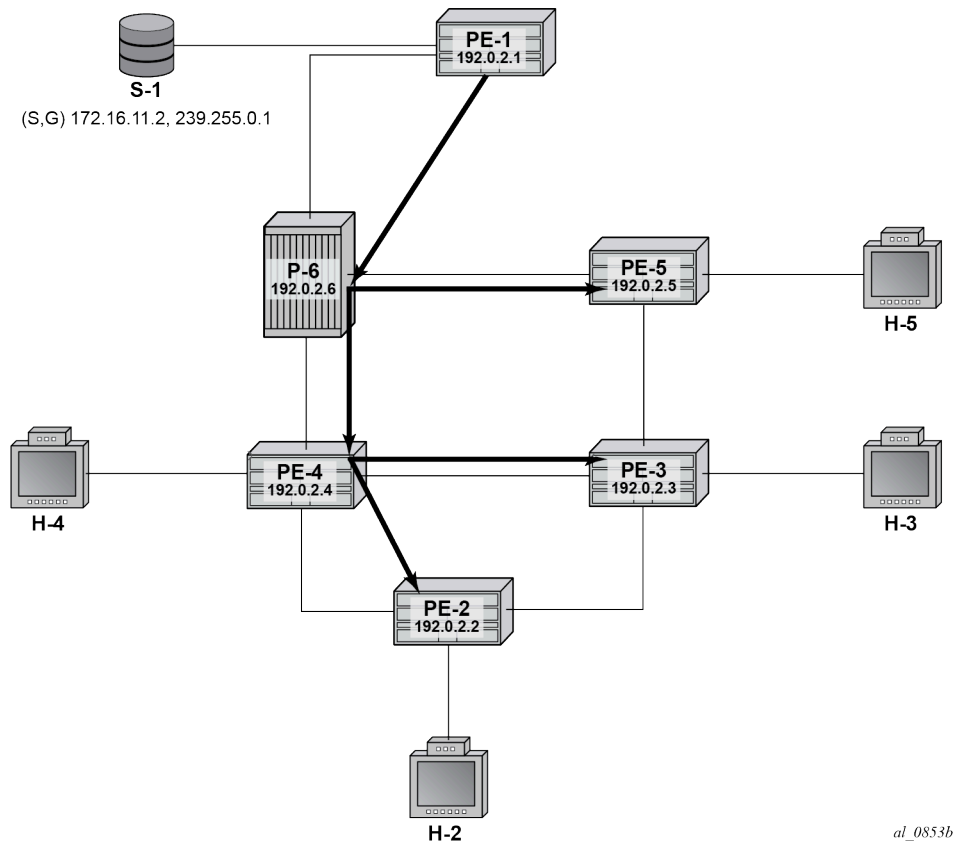
=====
Legend:  A = Active   S = Standby
=====
PIM Groups ipv4
=====
Group Address          Type          Spt Bit  Inc Intf  No.0ifs
Source Address        RP           State    Inc Intf(S)
-----
```

```

239.255.0.1      (*,G)          mpls-if-73729* 1
*                10.0.0.1
239.255.0.1      (S,G)          spt      mpls-if-73729* 1
172.16.11.2     10.0.0.1
-----
Groups : 2
=====
* indicates that the corresponding row element may have been truncated.
    
```

Assume that a receiver on each of PE-4 and PE-5 needs to join group 239.255.0.1, as shown in [Figure 63: S-PMSI P2MP LSP Schematic](#).

Figure 63: S-PMSI P2MP LSP Schematic



[Figure 63: S-PMSI P2MP LSP Schematic](#) shows the S-PMSI P2MP LSP. The next set of outputs shows the P2MP label mapping of the LDP LSP between PE-1 and the receiver PEs.

The root of the S-PMSI is on PE-1, as follows:

```

*A:PE-1# show router ldp bindings active p2mp p2mp-id 8193 root 192.0.2.1
=====
LDP Bindings (IPv4 LSR ID 192.0.2.1)
(IPv6 LSR ID ::)
=====
Label Status:
  U - Label In Use, N - Label Not In Use, W - Label Withdrawn
  WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
    
```

```

e - Label ELC
FEC Flags:
  LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
  BA - ASBR Backup FEC
=====
LDP Generic IPv4 P2MP Bindings (Active)
=====
P2MP-Id          Interface
RootAddr         Op
IngLbl           EgrLbl
EgrNH            EgrIf/LspId
-----
8193              73728
192.0.2.1        Push
--              524281
192.168.16.2     1/1/c1/1
-----
No. of Generic IPv4 P2MP Active Bindings: 1
=====

```

The egress label on PE-1 becomes the ingress label on P-6. P-6 has two leaves: one toward PE-4 and one toward PE-5, as follows:

```

*A:P-6# show router ldp bindings active p2mp p2mp-id 8193 root 192.0.2.1
=====
LDP Bindings (IPv4 LSR ID 192.0.2.6)
(IPv6 LSR ID ::)
=====
Label Status:
  U - Label In Use, N - Label Not In Use, W - Label Withdrawn
  WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
  e - Label ELC
FEC Flags:
  LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
  BA - ASBR Backup FEC
=====
LDP Generic IPv4 P2MP Bindings (Active)
=====
P2MP-Id          Interface
RootAddr         Op
IngLbl           EgrLbl
EgrNH            EgrIf/LspId
-----
8193              Unknw
192.0.2.1        Swap
524281           524280
192.168.46.1     1/1/c2/1

8193              Unknw
192.0.2.1        Swap
524281           524280
192.168.56.1     1/1/c1/1
-----
No. of Generic IPv4 P2MP Active Bindings: 2
=====

```

On PE-5, the following output shows that the LSP terminates as a leaf, as the operation (Op) is shown as "pop":

```
*A:PE-5# show router ldp bindings active p2mp p2mp-id 8193 root 192.0.2.1

=====
LDP Bindings (IPv4 LSR ID 192.0.2.5)
(IPv6 LSR ID ::)
=====
---snip---
=====
LDP Generic IPv4 P2MP Bindings (Active)
=====
P2MP-Id                Interface
RootAddr               Op
IngLbl                 EgrLbl
EgrNH                  EgrIf/LspId
-----
8193                   73729
192.0.2.1              Pop
524280                 --
--                     --

-----
No. of Generic IPv4 P2MP Active Bindings: 1
=====
```

On PE-4, the P2MP LSP has 3 entries: a pop operation to receiver H-4, and two label swaps toward PE-2 and PE-3:

```
*A:PE-4# show router ldp bindings active p2mp p2mp-id 8193 root 192.0.2.1

=====
LDP Bindings (IPv4 LSR ID 192.0.2.4)
(IPv6 LSR ID ::)
=====
Label Status:
  U - Label In Use, N - Label Not In Use, W - Label Withdrawn
  WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
  e - Label ELC
FEC Flags:
  LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
  BA - ASBR Backup FEC
=====
LDP Generic IPv4 P2MP Bindings (Active)
=====
P2MP-Id                Interface
RootAddr               Op
IngLbl                 EgrLbl
EgrNH                  EgrIf/LspId
-----
8193                   73729
192.0.2.1              Pop
524280                 --
--                     --

8193                   73729
192.0.2.1              Swap
524280                 524280
192.168.24.1          1/1/c2/1

8193                   73729
```

```
192.0.2.1          Swap
524280            524280
192.168.34.1     1/1/c3/1
```

```
-----
No. of Generic IPv4 P2MP Active Bindings: 3
=====
```

PE-2 and PE-3 are termination PEs for P2MP leaf. On PE-2, the pop operation is shown:

```
*A:PE-2# show router ldp bindings active p2mp p2mp-id 8193 root 192.0.2.1
```

```
=====
LDP Bindings (IPv4 LSR ID 192.0.2.2)
              (IPv6 LSR ID ::)
=====
```

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

```
=====
P2MP-Id      Interface
RootAddr     Op
IngLbl       EgrLbl
EgrNH        EgrIf/LspId
-----
8193         73729
192.0.2.1   Pop
524280      --
--          --
```

```
-----
No. of Generic IPv4 P2MP Active Bindings: 1
=====
```

On PE-3, the P2MP pop operation is shown:

```
*A:PE-3# show router ldp bindings active p2mp p2mp-id 8193 root 192.0.2.1
```

```
=====
LDP Bindings (IPv4 LSR ID 192.0.2.3)
              (IPv6 LSR ID ::)
=====
```

```
Label Status:
```

```
U - Label In Use, N - Label Not In Use, W - Label Withdrawn
WP - Label Withdraw Pending, BU - Alternate For Fast Re-Route
e - Label ELC
```

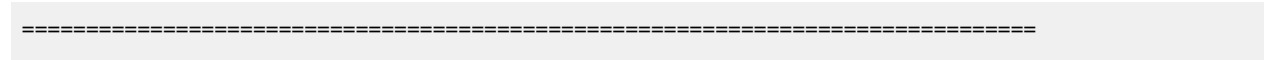
```
FEC Flags:
```

```
LF - Lower FEC, UF - Upper FEC, M - Community Mismatch,
BA - ASBR Backup FEC
```

```
=====
LDP Generic IPv4 P2MP Bindings (Active)
=====
```

```
P2MP-Id      Interface
RootAddr     Op
IngLbl       EgrLbl
EgrNH        EgrIf/LspId
-----
8193         73729
192.0.2.1   Pop
524280      --
--          --
```

```
-----
No. of Generic IPv4 P2MP Active Bindings: 1
=====
```



## Conclusion

MVPN wildcard Selective PMSI (S-PMSI), developed as per RFC 6625, provides an optimal solution for multicast routing in a VPRN. This protocol provides simple configuration, operation, and fast protection time in conjunction with MPLS and LDP fast-failover schemes. Wildcard S-PMSI can be used in a multicast network to avoid a large full mesh of an I-PMSI.

# Rosen MVPN Core Diversity

This chapter provides information about Rosen multicast virtual private network (MVPN) core diversity.

Topics in this chapter include:

- [Applicability](#)
- [Overview](#)
- [Configuration](#)
- [Conclusion](#)

## Applicability

This chapter was initially written for SR OS Release 13.0.R4, using Rosen MVPN. The CLI in the current edition is based on SR OS Release 23.10.R1.

Default multicast distribution trees (MDTs) for each virtual private routed network (VPRN) are signaled using protocol independent multicasting (PIM) and auto-discovery uses border gateway protocol multicast distribution tree sub-address family indicator (BGP MDT-SAFI) network layer routing.

## Overview

This chapter describes a service provider core network used by multiple content providers to deliver multicast services to multiple customers using Rosen MVPN. If the same set of PEs is used to deliver the MVPN, the MDTs are all routed across the same paths between the set of PEs. Because each MDT is signaled using PIM, and the source of all MDTs is the system address of the PE, the path to this source is the same.

Each remote PE then sends a PIM join toward this PE with its source address set to the system address. For multiple VPNs between the same set of PEs, the MDT follows the same path.

If there is a requirement to deliver content from each content provider across different MVPNs that use diversely routed MDTs, multiple IGP instances can be used: up to three different instances of IGP, OSPF or ISIS, can exist. In this chapter, two instances of OSPF are used to create incongruent topologies providing isolation between the MDTs of two different MVPNs: a default OSPF instance and OSPF instance 1. A separate /32 loopback address can be used as the MDT source address that is advertised in the non-default IGP, which can also be used as the BGP next hop for labeled IPv4 routes representing the customer source addresses.

Knowledge of multi-protocol BGP (MP-BGP) and RFC 4364, *BGP/MPLS IP VPNs*, is assumed throughout this chapter, as well as the original RFC 6037.

All PEs within an MVPN create a default MDT with their own system address as the source. Auto-discovery of PEs within a Rosen MVPN is achieved using the BGP route type of multicast data tree subsequent address family identifier (MDT-SAFI). Each PE originates an MDT-SAFI route update per MVPN. This route advertises the presence of the MVPN on a specific PE.

Each MDT-SAFI update contains attributes, including the following:

1. Route distinguisher
2. Route target extended community
3. MDT source address (usually the system address)
4. Group address of MDT

Upon receipt of an MDT-SAFI route update, each remote PE accepts or rejects the route based on the route target extended community value. If the route is accepted, a remote PE sends a PIM (S,G) join to this local PE. The (S,G) values are taken from the MDT-SAFI. The set of MDTs extend the c-multicast data tree across the MVPN and form PIM adjacencies between PEs within the MVPN. The neighbor address across the set of PIM-enabled tunnels is the default MDT source address, usually the system address.

When established, the default MDT is used to transport c-multicast PIM signaling between PEs.

If a source S, of a c-multicast group G, is connected to a sender PE, the route to the source is advertised to remote PEs as a BGP-labeled VPN-IPv4 route.

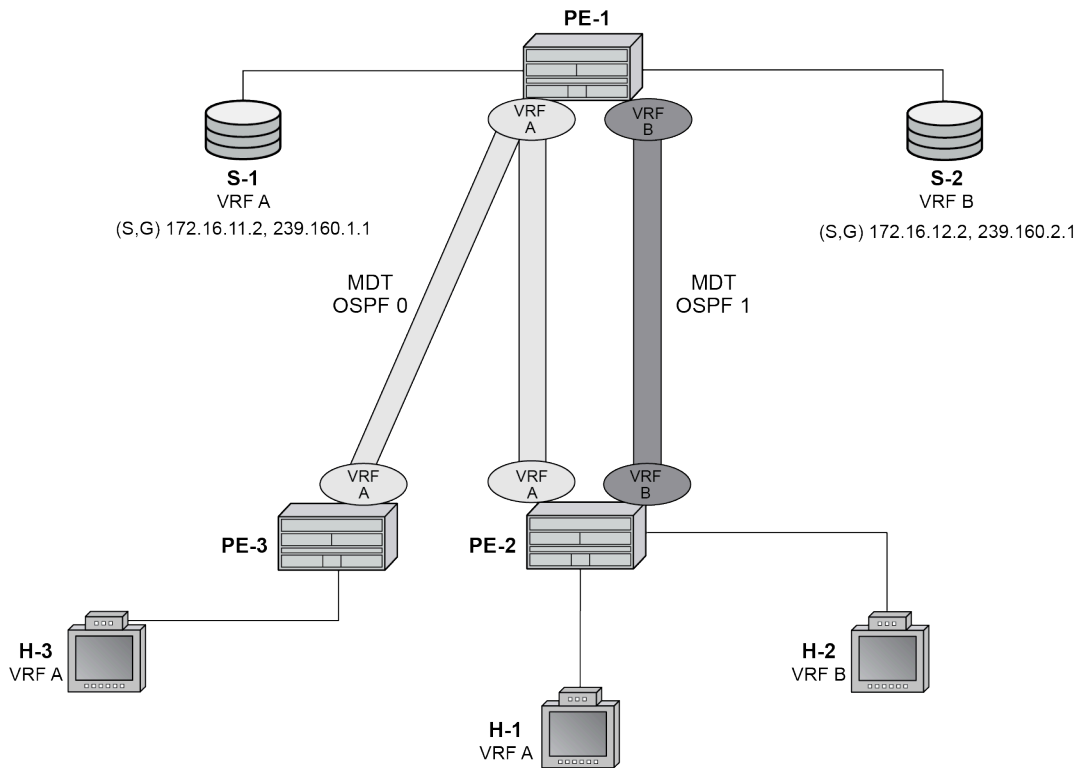
Therefore, an (S,G) join toward this source at a remote PE performs a reverse path forwarding (RPF) look-up of the unicast VRF table to find a suitable PIM-enabled interface. The next hop needs to be resolved to the MDT source address of the sender PE. A PIM join must now be forwarded toward the sender PE that has a PIM neighbor that matches the next hop for this route, the system address of the sender PE. This is the default MDT.

The system address is a significant address in this process. Any other VPRN that uses the same set of PEs also signals a set of default MDTs using a different group address, so they follow the same path between PEs across the provider network.

[Figure 64: Core Diversity Schematic](#) shows an example of core diversity; multicast sources provided by two separate content providers are connected to a provider network. There is a requirement to provide topology diversity so that the default MDTs between the same PEs are routed across different paths within the core.



Figure 64: Core Diversity Schematic



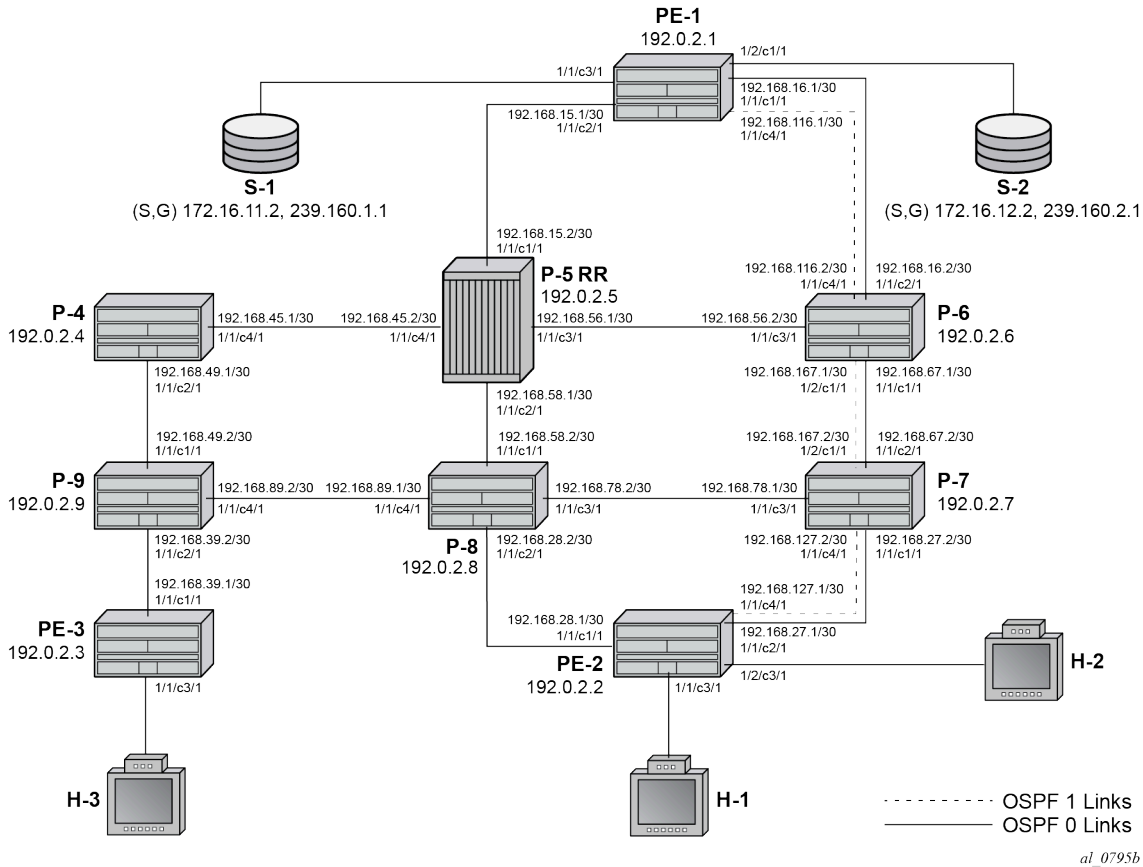
al\_0794b

Content servers from two separate content providers are connected to PE-1 with directly connected multicast sources. For simplicity, this example uses only a single multicast group for provider S-1 and S-2. Source S-1 is reachable via VRF A and source S-2 is reachable via VRF B.

Topology isolation for the multicast data trees of each VPRN can be provided by using two separate IGP instances; in this case, OSPF instances. Multi-instance IS-IS could also be used.

Figure 65: Core Diversity Network shows a schematic of the full network, including the c-multicast groups.

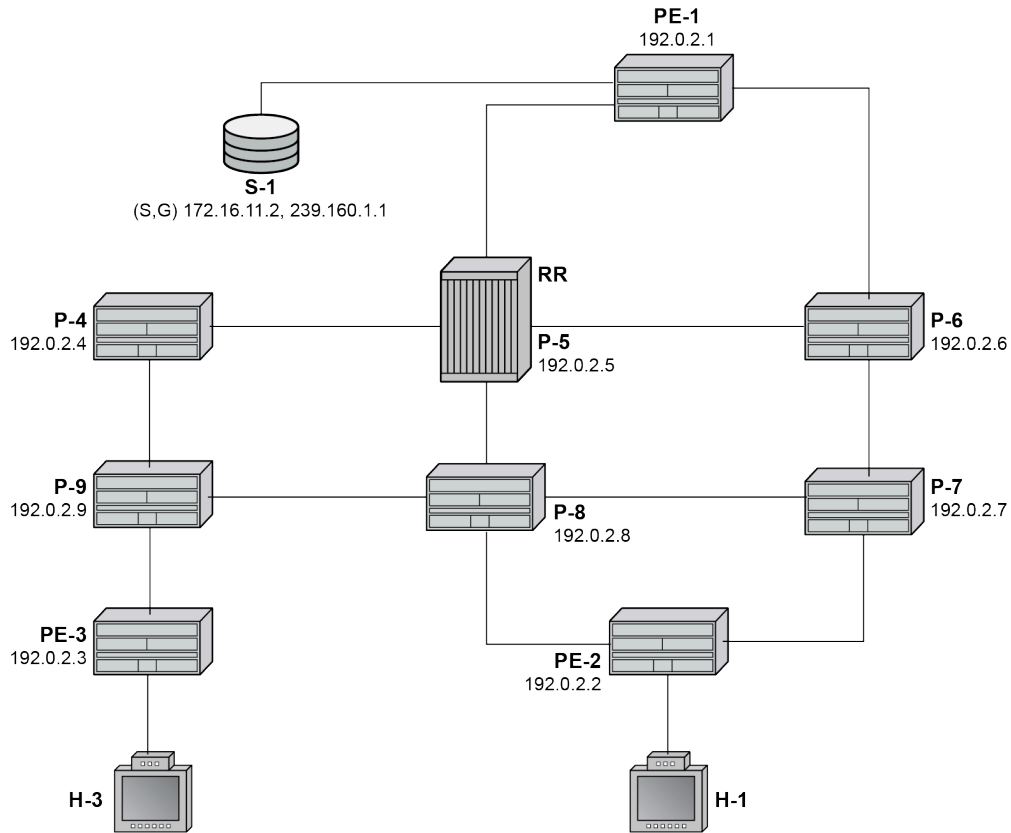
Figure 65: Core Diversity Network



All routers have interfaces in the OSPF base instance (instance 0) and routers interconnected by the dotted lines have interfaces in both the base instance and OSPF 1.

Figure 65: Core Diversity Network shows the extent of the OSPF base instance within the core network.

Figure 66: Core Diversity Network — Base OSPF

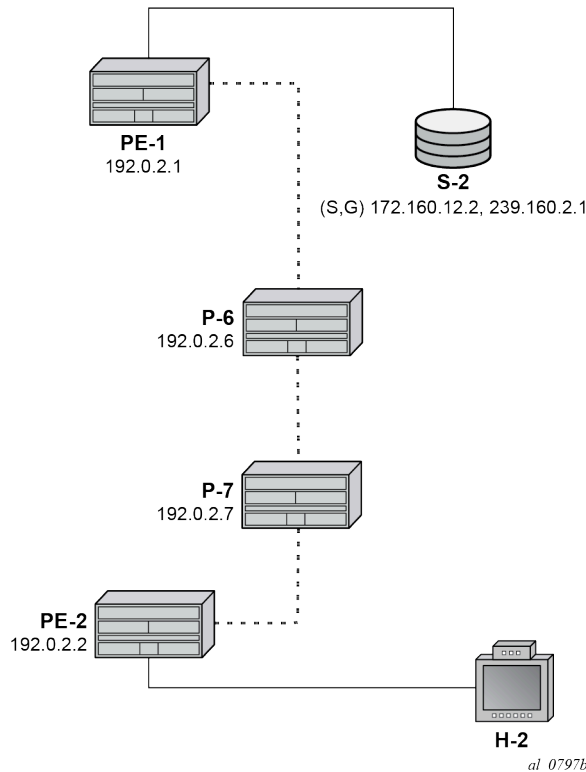


al\_0796b

In this case, assume that the shortest path between PE-1 and PE-2 is the path via P-5 and P-8.

Similarly, [Figure 67: Core Diversity Network - OSPF Instance 1](#) shows the extent of OSPF instance 1.

Figure 67: Core Diversity Network - OSPF Instance 1



The only path available between PE-1 and PE-2 is now completely diverse from the shortest path advertised between the same pair of PEs in the base OSPF instance.

Therefore, for any MDT to be signaled across the OSPF 1 topology, only addresses advertised within OSPF 1 must be used. As previously stated, the system address is used as the default MDT source address. This system address is not advertised within the OSPF 1 topology, so a replacement /32 loopback address is used as the default MDT source address within OSPF 1.

VPN-IPv4 routes that may represent a customer multicast source address should be reachable via the default MDT. In the non-default topology, the c-multicast signaling across the MVPN must resolve the c-multicast route via the MDT, which has its root at the non-default /32 loopback. Therefore, the VPN-IPv4 prefix representing the possible source routes needs to be advertised containing the non-default /32 loopback.

This can be achieved in one of two ways:

1. Use a route policy at the advertising PE that changes the BGP next hop to match the MDT source address for non-default topology MDTs.
2. Use the BGP connector attribute for all VPN-IPv4 route prefixes within a multicast VPRN that has auto-discovery set to MDT-SAFI. The connector attribute contains the MDT source address within the originator field.

This chapter describes the use of the connector attribute.

If the default IGP instance is used, the BGP next hop of the VPN-IPv4 route matches the source address of the default MDT.

Therefore, if a second /32 loopback is used that replaces the system address as MDT source address and also as the next hop for source address RPF look-up, the loopback could be advertised within the non-default IGP instance, and the paths between the PEs would follow this topology.

Core diversity is achieved by configuring the following steps:

1. Configuring multiple OSPF instances, as shown in [Figure 66: Core Diversity Network — Base OSPF](#) and [Figure 67: Core Diversity Network - OSPF Instance 1](#), and including the appropriate interfaces. This includes a separate loopback address per instance.
2. Configuring separate VPRNs with their own MDTs using BGP MDT-SAFI auto-discovery and PIM signaling across the appropriate PEs.
3. Configuring the VPRN that uses the base OSPF instance to use the system address as the source addresses for the MDTs (this is default behavior).
4. Configuring a separate loopback (/32) address that is advertised within OSPF instance 1 only.
5. Configuring the VPRN that uses the OSPF instance 1 to use the separate loopback system address as the source addresses for the MDTs.
6. Ensuring the unicast route that represents the c-source address is advertised as a VPN-IPv4 route and has a BGP connector attribute that contains an address that matches the MDT source address of the originating PE.

## Configuration

The following configuration tasks must be completed as a prerequisite:

- Full mesh OSPF base instance between each of the nodes. However, IS-IS could also be used for any or all of the IGP instances.
- Link-layer LDP between each P and PE router.
- PIM enabled on each router network interface.

## Global BGP Configuration

The first step is to configure an iBGP session between each of the PEs and the route reflector (P-5) shown in [Figure 65: Core Diversity Network](#). The address families negotiated between the iBGP peers are **vpn-ipv4**, for unicast routing, and **mdt-safi** for multicast routing.

The iBGP configuration for PE-1 is the following:

```
# on PE-1
configure
router
  autonomous-system 64496
  bgp
    group "INTERNAL"
      family vpn-ipv4 mdt-safi
      type internal
      neighbor 192.0.2.5
    exit
  exit
exit all
```

The configuration for the other PE nodes is the same.

P-5 is the route reflector for PE-1, PE-2, and PE-3, as follows:

```
# on P-5
configure
router
  autonomous-system 64496
  bgp
    cluster 0.0.0.1
    group "RR_CLIENTS"
      family vpn-ipv4 mdt-safi
      type internal
      neighbor 192.0.2.1
      exit
      neighbor 192.0.2.2
      exit
      neighbor 192.0.2.3
      exit
    exit
  exit
exit all
```

On PE-1, verify that the BGP session with the route reflector on P-5 is established with address families **mdt-safi** and **vpn-ipv4** capabilities negotiated:

```
*A:PE-1# show router bgp summary
=====
BGP Router ID:192.0.2.1      AS:64496      Local AS:64496
=====
BGP Admin State      : Up      BGP Oper State      : Up
Total Peer Groups    : 1      Total Peers         : 1
Total VPN Peer Groups : 0      Total VPN Peers     : 0
Current Internal Groups : 1      Max Internal Groups : 1
Total BGP Paths      : 26      Total Path Memory   : 9568
---snip---
=====
BGP Summary
=====
Legend : D - Dynamic Neighbor
=====
Neighbor
Description
          AS PktRcvd InQ  Up/Down  State|Rcv/Act/Sent (Addr Family)
          PktSent OutQ
-----
192.0.2.5
          64496      4    0 00h00m09s 0/0/0 (VpnIPv4)
          4    0    0/0/0 (MdtSafi)
-----
```

## Configuring VPRN on PEs

There are two VPRNs:

- VPRN 1 using the base instance OSPF topology. This is present on PE-1, PE-2, and PE-3.
- VPRN 2 using OSPF instance 1. This is present on PE-1 and PE-2.

The following output displays the configuration for VPRN 1 for the sender PE-1.

```
# on PE-1
configure
service
  vprn 1 name "VPRN 1" customer 1 create
  interface "int-PE-1-S-1" create
    address 172.16.11.1/24
    sap 1/1/c3/1 create
    exit
  exit
  bgp-ipvpn
  mpls
    auto-bind-tunnel
    resolution-filter
    ldp
    exit
    resolution filter
  exit
  route-distinguisher 64496:1
  vrf-target target:64496:1
  no shutdown
  exit
  exit
  pim
  apply-to all
  no shutdown
  exit
  mvpn
    auto-discovery mdt-safi
    provider-tunnel
    inclusive
    pim ssm 239.160.1.1
    exit
    exit
  vrf-target unicast
  exit
  exit
  no shutdown
  exit all
```

There is a single interface toward S-1, from which the multicast group is received.

PIM is enabled and applied to all interfaces.

The MVPN configuration enables BGP MDT-SAFI as the auto-discovery mechanism. The provider tunnels between the PEs within the MVPN are PIM SSM multicast data trees with a group address of 239.160.1.1.

The configuration for VPRN 1 for the receiver PE-2 is the following.

```
# on PE-2
configure
service
  vprn 1 name "VPRN 1" customer 1 create
  interface "int-PE-2-H-1" create
    address 172.16.21.1/24
    sap 1/1/c3/1 create
    exit
  exit
  bgp-ipvpn
  mpls
    auto-bind-tunnel
```

```

        resolution-filter
        ldp
        exit
        resolution filter
    exit
    route-distinguisher 64496:1
    vrf-target target:64496:1
    no shutdown
exit
exit
igmp
interface "int-PE-2-H-1"
    no shutdown
exit
no shutdown
exit
pim
    apply-to all
    no shutdown
exit
mvpn
    auto-discovery mdt-safi
    provider-tunnel
        inclusive
        pim ssm 239.160.1.1
        exit
    exit
    vrf-target unicast
    exit
exit
no shutdown
exit all

```

The configuration for VPRN 1 for receiver PE-3 is as follows.

```

# on PE-3
configure
  service
    vprn 1 name "VPRN 1" customer 1 create
    interface "int-PE-3-H-3" create
      address 172.16.31.1/24
      sap 1/1/c3/1 create
      exit
    exit
    bgp-ipvpn
      mpls
        auto-bind-tunnel
        resolution-filter
        ldp
        exit
        resolution filter
      exit
      route-distinguisher 64496:1
      vrf-target target:64496:1
      no shutdown
    exit
  exit
  igmp
    interface "int-PE-3-H-3"
      no shutdown
    exit
  no shutdown

```



```

exit
pim
  apply-to all
  no shutdown
exit
mvpn
  auto-discovery mdt-safi
  provider-tunnel
  inclusive
  pim ssm 239.160.1.1
  exit
  exit
  vrf-target unicast
  exit
  no shutdown
exit all

```

At PE-2, the MDT SAFI NLRI advertised by PE-1 is as follows:

```
*A:PE-2# show router bgp routes mdt-safi grp-address 239.160.1.1 source-ip 192.0.2.1 detail
```

```

=====
BGP Router ID:192.0.2.2      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP MDT-SAFI Routes
=====
Original Attributes

Route Dist.   : 64496:1
Source Addr   : 192.0.2.1
Group Addr    : 239.160.1.1
Nextthop     : 192.0.2.1
From          : 192.0.2.5
Res. Nextthop : 0.0.0.0
Local Pref.   : 100
Aggregator AS : None
Atomic Aggr.  : Not Atomic
AIGP Metric   : None
Connector     : None
Community     : target:64496:1
Cluster       : 0.0.0.1
Originator Id : 192.0.2.1
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
Source Class  : 0
Add Paths Send : Default
Last Modified : 00h01m09s

Interface Name : NotAvailable
Aggregator     : None
MED            : 0
IGP Cost       : n/a

Peer Router Id : 192.0.2.5
Final Orig Val : N/A
Dest Class     : 0

Modified Attributes
---snip---
-----

```

```
Routes : 1
=====
```

The source and group address is used by PE-2 (and PE-3) to join the MDT that has its root on PE-1. The source address used is the system address of PE-1.

Examining the MDTs for this VPRN on PE-1 shows the state, as follows:

```
*A:PE-1# show router pim group 239.160.1.1
=====
Legend:  A = Active   S = Standby
=====
PIM Groups ipv4
=====
Group Address          Type          Spt Bit   Inc Intf   No.0ifs
  Source Address       RP
-----
239.160.1.1           (S,G)         spt       system     2
  192.0.2.1
239.160.1.1           (S,G)         spt       int-PE-1-P-5  1
  192.0.2.2
239.160.1.1           (S,G)         spt       int-PE-1-P-5  1
  192.0.2.3
-----
Groups : 3
=====
```

The MDT with the root of its tree on PE-1 is as follows:

```
*A:PE-1# show router pim group 239.160.1.1 detail source 192.0.2.1
=====
PIM Source Group ipv4
=====
Group Address       : 239.160.1.1
Source Address      : 192.0.2.1
RP Address          : 0
Advt Router         : 192.0.2.1
Flags               : spt                Type           : (S,G)
Mode                : sparse
MRIB Next Hop      :
MRIB Src Flags     : self
Keepalive Timer Exp: 0d 00:03:14
Up Time            : 0d 00:01:49      Resolved By     : rtable-u

Up JP State         : Joined           Up JP Expiry    : 0d 00:00:10
Up JP Rpt           : Not Joined StarG  Up JP Rpt Override : 0d 00:00:00

Register State     : No Info
Reg From Anycast RP: No

Rpf Neighbor       :
Incoming Intf      : system
Outgoing Intf List : system, int-PE-1-P-5

Curr Fwding Rate   : 0.000 kbps
Forwarded Packets  : 8                Discarded Packets : 0
Forwarded Octets   : 624              RPF Mismatches    : 0
Spt threshold      : 0 kbps           ECMP opt threshold : 7
Admin bandwidth    : 1 kbps
-----
Groups : 1
```

The source address of the tree is the system address of the router, which is determined from the MDT SAFI NLRI that is advertised to all other PEs via the route reflector. Also, the outgoing interface list contains an interface (int-PE-1-P-5) that is OSPF enabled, and advertised within the base OSPF instance.

From the MDT on PE-2, which has its root on PE-1, the incoming interface is an OSPF interface advertised in the base OSPF instance, as shown.

```
*A:PE-2# show router pim group 239.160.1.1 detail source 192.0.2.1
=====
PIM Source Group ipv4
=====
Group Address      : 239.160.1.1
Source Address     : 192.0.2.1
RP Address         : 0
Advt Router       : 192.0.2.1
Flags              : spt                Type           : (S,G)
Mode               : sparse
MRIB Next Hop     : 192.168.28.2
MRIB Src Flags    : remote
Keepalive Timer Exp: 0d 00:03:12
Up Time           : 0d 00:01:38         Resolved By      : rtable-u

Up JP State       : Joined              Up JP Expiry     : 0d 00:00:22
Up JP Rpt        : Not Joined StarG    Up JP Rpt Override : 0d 00:00:00

Register State    : No Info
Reg From Anycast RP: No

Rpf Neighbor      : 192.168.28.2
Incoming Intf   : int-PE-2-P-8
Outgoing Intf List : system

Curr Fwding Rate  : 0.000 kbps
Forwarded Packets : 4                  Discarded Packets : 0
Forwarded Octets  : 312                 RPF Mismatches    : 0
Spt threshold     : 0 kbps              ECMP opt threshold : 7
Admin bandwidth   : 1 kbps

-----
Groups : 1
=====
```

The incoming interface shown is "int-PE-2-P-8". Similarly for PE-3, the incoming interface is "int-PE-3-P-9".

```
*A:PE-3# show router pim group 239.160.1.1 detail source 192.0.2.1
=====
PIM Source Group ipv4
=====
Group Address      : 239.160.1.1
Source Address     : 192.0.2.1
RP Address         : 0
Advt Router       : 192.0.2.1
Flags              : spt                Type           : (S,G)
Mode               : sparse
MRIB Next Hop     : 192.168.39.2
MRIB Src Flags    : remote
Keepalive Timer Exp: 0d 00:03:09
Up Time           : 0d 00:01:29         Resolved By      : rtable-u
```

```

Up JP State      : Joined           Up JP Expiry      : 0d 00:00:30
Up JP Rpt       : Not Joined StarG  Up JP Rpt Override : 0d 00:00:00

Register State   : No Info
Reg From Anycast RP: No

Rpf Neighbor     : 192.168.39.2
Incoming Intf  : int-PE-3-P-9
Outgoing Intf List : system

Curr Fwding Rate : 0.000 kbps
Forwarded Packets : 3                Discarded Packets : 0
Forwarded Octets  : 234              RPF Mismatches    : 0
Spt threshold    : 0 kbps            ECMP opt threshold : 7
Admin bandwidth  : 1 kbps

-----
Groups : 1
=====

```

## VPRN Using Non-Default IGP Instance

A VPRN instance is configured on each of PE-1 and PE-2 that uses an MDT topology governed by the non-default OSPF instance.

Additional interfaces need to be configured.

```

# on PE-1
configure
router
  interface "int-PE-1-P-6a"
    address 192.168.116.1/30
    port 1/1/c4/1
  exit
  interface "loop-1"
    loopback
    address 192.0.3.1/32
  exit
exit all

```

There are parallel links between PE-1 and P-6. The interface name of the second link contains the suffix *a*.

In a Rosen MVPN, each PE constructs a default MDT to all other PEs in the multicast VPN domain, as defined by the MDT SAFI BGP update received. The MDT update contains the source address of the MDT to which each PE should join.

When each of the other PEs receives the MDT SAFI network layer reachability information (NLRI), a PIM join is sent to the source address within the global PIM routing instance to create the MDT.

The MDT source address is usually the system address. Because the system address is advertised in the base instance of OSPF, another /32 address must be used as the source address for the default MDT. Therefore, a second loopback address is configured and used as the source address for the default MDT. For PE-1, this interface is called loop-1, and it is advertised in OSPF 1.

The interface loop-1 is used as the source address for the MDTs and the next hop for the unicast route representing the source address of the c-multicast group.

The non-default OSPF instance for PE-1 is configured as follows, where 192.0.3.1 is the OSPF 1 router-ID. The router ID need not be equal to the IP address for loop-1, but in this case it is.

```
# on PE-1
configure
router
  ospf 1 192.0.3.1
    area 0.0.0.0
      interface "loop-1"
      exit
      interface "int-PE-1-P-6a"
        interface-type point-to-point
      exit
    exit
  no shutdown
exit all
```

LDP is also required for BGP next-hop resolution and is configured as follows for PE-1.

```
# on PE-1
configure
router
  ldp
    interface-parameters
      interface "int-PE-1-P-6a"
        ipv4
          transport-address interface
      exit all
```

The transport address is set to interface, rather than the default of system address; this is because the system address is not reachable within OSPF instance 1.

For completeness, the configuration of the additional interfaces, OSPF instance 1 and LDP of PE-2 is shown in the following three outputs.

```
# on PE-2
configure
router
  interface "int-PE-2-P-7a"
    address 192.168.127.1/30
    port 1/1/c4/1
  exit
  interface "loop-1"
    loopback
    address 192.0.3.2/32
  exit
exit all
```

The OSPF 1 instance configuration is as follows:

```
# on PE-2
configure
router
  ospf 1 192.0.3.2
    area 0.0.0.0
      interface "loop-1"
      exit
      interface "int-PE-2-P-7a"
        interface-type point-to-point
      exit
    exit
```

```
no shutdown
exit all
```

The LDP configuration is as follows:

```
# on PE-2
configure
router
  ldp
    interface-parameters
      interface "int-PE-2-P-7a"
        ipv4
          transport-address interface
        exit all
```

PIM needs to be enabled on all interfaces.

The MDT source address for VPRN 2 is the loop-1 address. Each PE within this VPRN has to join the MDT sourced on PE-1, so the MDT SAFI NLRI must advertise the source address of the MDT group as loop-1. This is achieved by specifying the MDT SAFI source address within the MVPN context. The following output displays the VPRN configuration for PE-1.

```
# on PE-1
configure
service
  vprn 2 name "VPRN 2" customer 1 create
  interface "int-PE-1-S-2" create
    address 172.16.12.1/24
    sap 1/2/cl/1 create
    exit
  exit
  bgp-ipvpn
  mpls
    auto-bind-tunnel
    resolution-filter
    ldp
    exit
    resolution filter
  exit
  route-distinguisher 64496:2
  vrf-target target:64496:2
  no shutdown
  exit
  exit
  pim
    apply-to all
    no shutdown
  exit
  mvpn
    auto-discovery mdt-safi source-address 192.0.3.1
    provider-tunnel
    inclusive
    pim ssm 239.160.2.1
    exit
    exit
  exit
  vrf-target target:64496:2
  exit
  exit
  no shutdown
exit all
```

The MDT SAFI source address modification is only required on PEs that use the non-default /32 addresses. The system address must not be explicitly configured as the MDT source address for MVPNs that use the default IGP instance. As previously stated, only three MVPNs can be used to create core diversity, one of which must be the default instance. Configuring the system address as a source address prevents the creation of a third MVPN because only two MVPNs are allowed to use explicitly configured MDT source addresses.

## Verification of Core Diversity

The MDT SAFI NLRI advertised by the PE-1 sender router contains the following information.

```
*A:PE-1# show router bgp routes mdt-safi hunt rd 64496:2 | match "RIB Out" post-lines 25 pre-
lines 1
-----
RIB Out Entries
-----
Route Dist.      : 64496:2
Source Addr    : 192.0.3.1
Group Addr      : 239.160.2.1
Nexthop         : 192.0.2.1
To              : 192.0.2.5
Res. Nexthop    : n/a
Local Pref.     : 100
Aggregator AS   : None
Atomic Aggr.    : Not Atomic
AIGP Metric     : None
Connector       : None
Community       : target:64496:2
Cluster         : No Cluster Members
Originator Id   : None
Origin          : IGP
AS-Path         : No As-Path
Route Tag       : 0
Neighbor-AS     : n/a
DB Orig Val     : N/A
Source Class    : 0
Interface Name  : NotAvailable
Aggregator     : None
MED            : 0
IGP Cost       : n/a
Peer Router Id  : 192.0.2.5
Final Orig Val  : N/A
Dest Class     : 0
-----
Routes : 3
=====
```

The source address is set to 192.0.3.1, which is the address of the loopback address used in the non-default OSPF instance 1 of PE-1.

The following output shows the MDT that has its root on PE-1, and that the source address is set to 192.0.3.1. The outgoing interface list includes the router interface contained within the OSPF 1 instance, proving that the non-default OSPF instance is used.

```
*A:PE-1# show router pim group 239.160.2.1 source 192.0.3.1 detail
=====
PIM Source Group ipv4
=====
Group Address    : 239.160.2.1
Source Address : 192.0.3.1
RP Address       : 0
Advt Router    : 192.0.2.1
Flags           : spt
Mode            : sparse
Type            : (S,G)
```

```

MRIB Next Hop      :
MRIB Src Flags     : self
Keepalive Timer Exp: 0d 00:03:28
Up Time           : 0d 00:00:36      Resolved By      : rtable-u

Up JP State        : Joined           Up JP Expiry      : 0d 00:00:24
Up JP Rpt          : Not Joined StarG Up JP Rpt Override: 0d 00:00:00

Register State     : No Info
Reg From Anycast RP: No

Rpf Neighbor       :
Incoming Intf      : loop-1
Outgoing Intf List : system, int-PE-1-P-6a

Curr Fwding Rate   : 0.000 kbps
Forwarded Packets  : 5                Discarded Packets : 0
Forwarded Octets   : 390              RPF Mismatches    : 0
Spt threshold      : 0 kbps           ECMP opt threshold: 7
Admin bandwidth    : 1 kbps
-----
Groups : 1
=====

```

The PIM interfaces within VPRN 2 are now present on PE-1, as follows:

```

*A:PE-1# show router 2 pim interface

=====
PIM Interfaces ipv4
=====
Interface          Adm  Opr  DR Prty      Hello Intvl  Mcast Send
DR
-----
int-PE-1-S-2       Up   Up   1           30          auto
  172.16.12.1
2-mt-239.160.2.1   Up   Up   1           N/A         auto
  192.0.3.2
-----
Interfaces : 2 Tunnel-Interfaces : 0
=====

```

Likewise, for PE-2, the PIM interfaces within VPRN 2 are displayed, as follows:

```

*A:PE-2# show router 2 pim interface

=====
PIM Interfaces ipv4
=====
Interface          Adm  Opr  DR Prty      Hello Intvl  Mcast Send
DR
-----
int-PE-2-H-2       Up   Up   1           30          auto
  172.16.22.1
2-mt-239.160.2.1   Up   Up   1           N/A         auto
  192.0.3.2
-----
Interfaces : 2 Tunnel-Interfaces : 0
=====

```



Within the VPRN, there are PIM neighbors shown via the MDT. On PE-2, the PIM neighbor is 192.0.3.1, as follows:

```
*A:PE-2# show router 2 pim neighbor
=====
PIM Neighbor ipv4
=====
Interface          Nbr DR Prty    Up Time      Expiry Time  Hold Time
  Nbr Address
-----
2-mt-239.160.2.1   1              0d 00:00:22  0d 00:01:23  105
  192.0.3.1
-----
Neighbors : 1
=====
```

The PIM interface on PE-2 designated as 2-mt-239.160.2.1 with a neighbor address of 192.0.3.1 is the MDT interface toward PE-1.

The prefix that represents the source address on PE-1 is advertised as a VPN-IPv4 route, which contains a BGP connector attribute.

This can be shown when the VPN-IPv4 route is examined on PE-2, as follows:

```
*A:PE-2# show router bgp routes 172.16.12.0/24 vpn-ipv4 hunt
=====
BGP Router ID:192.0.2.2      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP VPN-IPv4 Routes
=====
-----
RIB In Entries
-----
Network       : 172.16.12.0/24
Nextthop      : 192.0.2.1
Route Dist.   : 64496:2          VPN Label     : 524277
Path Id       : None
From          : 192.0.2.5
Res. Nextthop : n/a
Local Pref.   : 100
Aggregator AS : None           Interface Name : int-PE-2-P-8
Atomic Aggr.  : Not Atomic    Aggregator    : None
AIGP Metric   : None          MED           : None
Connector     : RD 64496:2, Originator 192.0.3.1
Community     : target:64496:2
Cluster       : 0.0.0.1
Originator Id : 192.0.2.1          Peer Router Id : 192.0.2.5
Fwd Class     : None          Priority       : None
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  : 00h00m41s
VPRN Imported  : 2
```

```
-----
RIB Out Entries
-----
```

```
Routes : 1
=====
```

The originator value within the connector attribute is shown to be 192.0.3.1, which is the same as the MDT source address of PE-1. The BGP next hop is still set to the system address of PE-1, so the unicast route can still be resolved via an LDP tunnel.

PIM now resolves the c-source address RPF using the originator value within the connector attribute.

Similarly, for VPRN 1, the route on PE-1 representing the source address is also advertised as a VPN-IPv4 address that contains a BGP connector attribute.

```
*A:PE-2# show router bgp routes 172.16.11.0/24 vpn-ipv4 hunt
```

```
=====
BGP Router ID:192.0.2.2      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
```

```
Origin codes  : i - IGP, e - EGP, ? - incomplete
```

```
=====
BGP VPN-IPv4 Routes
=====
```

```
-----
RIB In Entries
-----
```

```
Network       : 172.16.11.0/24
NextHop       : 192.0.2.1
Route Dist.   : 64496:1          VPN Label     : 524278
Path Id       : None
From          : 192.0.2.5
Res. NextHop  : n/a
Local Pref.   : 100
Aggregator AS : None           Interface Name : int-PE-2-P-8
Atomic Aggr.  : Not Atomic     Aggregator    : None
AIGP Metric   : None           MED           : None
Connector     : RD 64496:1, Originator 192.0.2.1
Community     : target:64496:1
Cluster       : 0.0.0.1
Originator Id : 192.0.2.1       Peer Router Id : 192.0.2.5
Fwd Class     : None           Priority       : None
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  : 00h00m41s
VPRN Imported  : 1
-----
```

RIB Out Entries

```
-----
-----
Routes : 1
=====
```

## Verification of Multicast Traffic

An IGMPv3 query is initiated from all 3 hosts: H-1, H-2, and H-3 in Figure 1, and the multicast streams from S-1 and S-2 into interfaces on the two VPRNs are enabled.

Consider VPRN 1, which uses the default topology. On PE-1, the group 239.160.1.123 can be shown. The outgoing and incoming interface lists are populated, with the outgoing interface being the MDT interface for the VPRN:

```
*A:PE-1# show router 1 pim group detail
=====
PIM Source Group ipv4
=====
Group Address      : 239.160.1.123
Source Address     : 172.16.11.2
RP Address         : 0
Advt Router        : 192.0.2.1
Flags              :                               Type           : (S,G)
Mode               : sparse
MRIB Next Hop      : 172.16.11.2
MRIB Src Flags     : direct
Keepalive Timer    : Not Running
Up Time            : 0d 00:01:04      Resolved By         : rtable-u

Up JP State        : Joined           Up JP Expiry        : 0d 00:00:00
Up JP Rpt          : Not Joined StarG  Up JP Rpt Override  : 0d 00:00:00

Register State     : No Info
Reg From Anycast RP: No

Rpf Neighbor       : 172.16.11.2
Incoming Intf    : int-PE-1-S-1
Outgoing Intf List : 1-mt-239.160.1.1

Curr Fwding Rate   : 9627.528 kbps
Forwarded Packets  : 78934             Discarded Packets   : 0
Forwarded Octets   : 77513188         RPF Mismatches      : 0
Spt threshold      : 0 kbps           ECMP opt threshold  : 7
Admin bandwidth    : 1 kbps

-----
Groups : 1
=====
```

The same groups can be shown within VPRN 1 on PE-2.

```
*A:PE-2# show router 1 pim group detail
=====
PIM Source Group ipv4
=====
Group Address      : 239.160.1.123
Source Address     : 172.16.11.2
RP Address         : 0
```

```

Advt Router      : 192.0.2.1
Flags            :                               Type           : (S,G)
Mode            : sparse
MRIB Next Hop   : 192.0.2.1
MRIB Src Flags  : remote
Keepalive Timer : Not Running
Up Time         : 0d 00:00:38      Resolved By        : rtable-u

Up JP State     : Joined           Up JP Expiry       : 0d 00:01:06
Up JP Rpt      : Not Joined StarG Up JP Rpt Override : 0d 00:00:00

Register State  : No Info
Reg From Anycast RP: No

Rpf Neighbor    : 192.0.2.1
Incoming Intf : 1-mt-239.160.1.1
Outgoing Intf List : int-PE-2-H-1

Curr Fwding Rate : 4737.168 kbps
Forwarded Packets : 33582           Discarded Packets : 0
Forwarded Octets  : 32977524      RPF Mismatches    : 0
Spt threshold    : 0 kbps         ECMP opt threshold : 7
Admin bandwidth  : 1 kbps
-----
Groups : 1
=====

```

The MDT is now the incoming interface with an upstream RPF neighbor of 192.0.2.1, the system address of PE-1. Similarly for PE-3:

```

*A:PE-3# show router 1 pim group detail

=====
PIM Source Group ipv4
=====
Group Address    : 239.160.1.123
Source Address   : 172.16.11.2
RP Address       : 0
Advt Router      : 192.0.2.1
Flags            :                               Type           : (S,G)
Mode            : sparse
MRIB Next Hop   : 192.0.2.1
MRIB Src Flags  : remote
Keepalive Timer : Not Running
Up Time         : 0d 00:01:10      Resolved By        : rtable-u

Up JP State     : Joined           Up JP Expiry       : 0d 00:00:51
Up JP Rpt      : Not Joined StarG Up JP Rpt Override : 0d 00:00:00

Register State  : No Info
Reg From Anycast RP: No

Rpf Neighbor    : 192.0.2.1
Incoming Intf : 1-mt-239.160.1.1
Outgoing Intf List : int-PE-3-H-3

Curr Fwding Rate : 4737.168 kbps
Forwarded Packets : 5305           Discarded Packets : 0
Forwarded Octets  : 5209510      RPF Mismatches    : 0
Spt threshold    : 0 kbps         ECMP opt threshold : 7
Admin bandwidth  : 1 kbps
-----
Groups : 1

```

Consider VPRN 2, which uses the non-default topology. On PE-1, the group 239.160.2.123 can be shown. The outgoing and incoming interface lists are populated, with the outgoing interface being the MDT interface for the VPRN.

```

=====
*A:PE-1# show router 2 pim group detail
=====
PIM Source Group ipv4
=====
Group Address      : 239.160.2.123
Source Address     : 172.16.12.2
RP Address         : 0
Advt Router        : 192.0.2.1
Flags              :                               Type           : (S,G)
Mode               : sparse
MRIB Next Hop     : 172.16.12.2
MRIB Src Flags    : direct
Keepalive Timer   : Not Running
Up Time           : 0d 00:00:51      Resolved By        : rtable-u

Up JP State       : Joined           Up JP Expiry       : 0d 00:00:00
Up JP Rpt        : Not Joined StarG  Up JP Rpt Override : 0d 00:00:00

Register State    : No Info
Reg From Anycast RP: No

Rpf Neighbor      : 172.16.12.2
Incoming Intf   : int-PE-1-S-2
Outgoing Intf List : 2-mt-239.160.2.1

Curr Fwding Rate  : 9623.600 kbps
Forwarded Packets : 62512           Discarded Packets  : 0
Forwarded Octets  : 61386784      RPF Mismatches    : 0
Spt threshold     : 0 kbps         ECMP opt threshold : 7
Admin bandwidth   : 1 kbps

-----
Groups : 1
=====

```

The outgoing interface list is again populated with the MDT being the interface. This MDT is encapsulated in the multicast tree shown in the global PIM context as multicast group 239.160.2.1 with source address 192.0.3.1. This can be shown to have an outgoing interface list containing the interface int-PE-1-P-6a, which is an OSPF 1 interface and was shown in a preceding output.

```

=====
*A:PE-1# show router pim group detail 239.160.2.1
=====
PIM Source Group ipv4
=====
Group Address      : 239.160.2.1
Source Address     : 192.0.3.1
RP Address         : 0
Advt Router        : 192.0.2.1
Flags              : spt                               Type           : (S,G)
Mode               : sparse
MRIB Next Hop     :
MRIB Src Flags    : self
Keepalive Timer Exp: 0d 00:03:04
Up Time           : 0d 00:07:00      Resolved By        : rtable-u

```

```

Up JP State      : Joined           Up JP Expiry      : 0d 00:00:00
Up JP Rpt       : Not Joined StarG  Up JP Rpt Override : 0d 00:00:00

Register State   : No Info
Reg From Anycast RP: No

Rpf Neighbor    :
Incoming Intf   : loop-1
Outgoing Intf List : system, int-PE-1-P-6a

Curr Fwding Rate : 9627.885 kbps
Forwarded Packets : 254784           Discarded Packets : 0
Forwarded Octets  : 250182520       RPF Mismatches    : 0
Spt threshold    : 0 kbps           ECMP opt threshold : 7
Admin bandwidth  : 1 kbps
    
```

```

=====
PIM Source Group ipv4
=====
    
```

```

Group Address    : 239.160.2.1
Source Address   : 192.0.3.2
RP Address       : 0
Advt Router     : 192.0.3.2
Flags           : spt                Type           : (S,G)
Mode            : sparse
MRIB Next Hop   : 192.168.116.2
MRIB Src Flags  : remote
Keepalive Timer Exp: 0d 00:03:16
Up Time         : 0d 00:06:49       Resolved By      : rtable-u

Up JP State      : Joined           Up JP Expiry      : 0d 00:00:11
Up JP Rpt       : Not Joined StarG  Up JP Rpt Override : 0d 00:00:00

Register State   : No Info
Reg From Anycast RP: No

Rpf Neighbor    : 192.168.116.2
Incoming Intf   : int-PE-1-P-6a
Outgoing Intf List : system

Curr Fwding Rate : 0.000 kbps
Forwarded Packets : 21             Discarded Packets : 0
Forwarded Octets  : 1638           RPF Mismatches    : 0
Spt threshold    : 0 kbps           ECMP opt threshold : 7
Admin bandwidth  : 1 kbps
    
```

```

-----
Groups : 2
=====
    
```

## Conclusion

MVPN Core Diversity allows service providers to provide separation in terms of topology between content providers that use a core network to provide transport between source and receivers in a multicast VPN. This chapter provides the configuration for multiple instances of OSPF which, together with the associated commands and outputs, can be used for verifying and troubleshooting.

# Rosen MVPN Inter-AS Option B

This chapter provides information about Rosen MVPN: Inter-AS Option B configurations.

Topics in this chapter include:

- [Applicability](#)
- [Overview](#)
- [Configuration](#)
- [Conclusion](#)

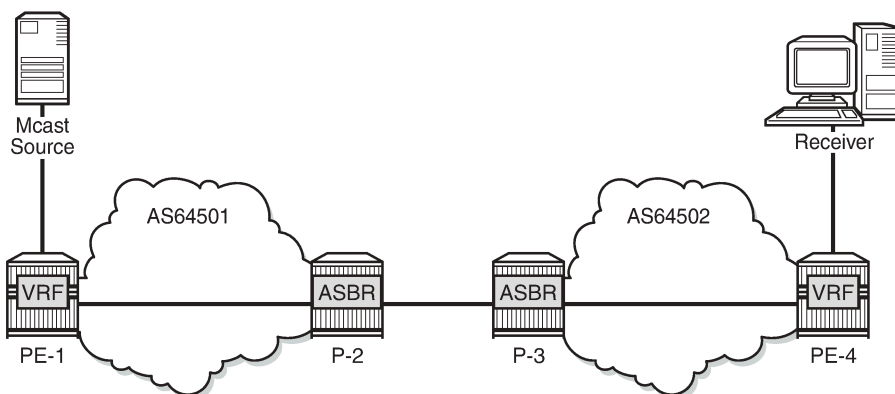
## Applicability

This chapter was initially written for SR OS Release 11.0.R3. The CLI in the current edition is based on SR OS Release 23.10.R1. Knowledge of the Nokia multicast and Layer 3 VPN concepts are assumed throughout this document.

## Overview

This chapter covers a basic technology overview, the network topology and configuration examples which are used for multicast virtual private network (MVPN) inter-autonomous system (AS) option B. The Inter-AS MVPN feature allows the setup of multicast distribution trees (MDTs) spanning multiple autonomous systems.

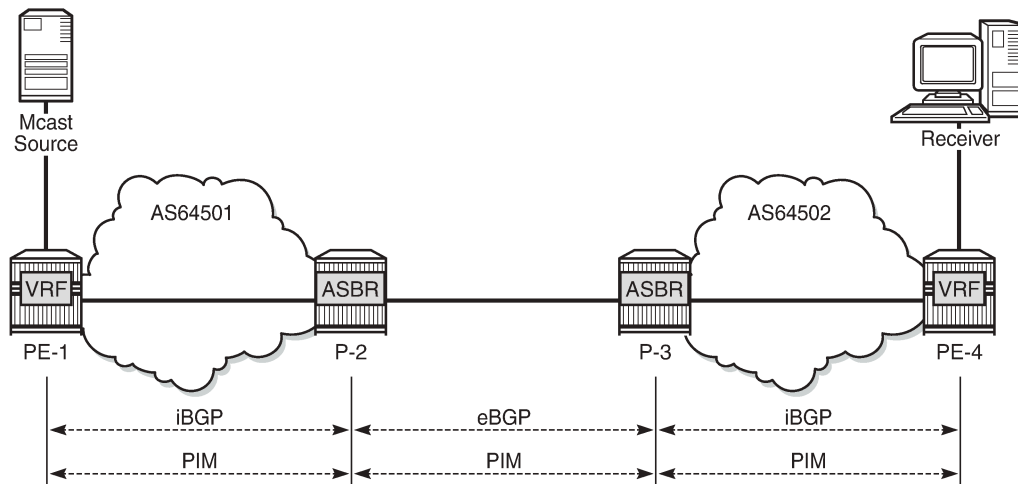
Figure 68: General Topology for Inter-AS MVPN



25402

This chapter covers Rosen MVPN Inter-AS support (Option-B). Inter-AS Option B is supported for protocol independent multicast (PIM) source-specific multicast (SSM) with Rosen MVPN using multicast distribution tree (MDT) subsequent address family indicator (SAFI), the border gateway protocol (BGP) connector attribute and PIM reverse path forwarding (RPF) vector.

Figure 69: Protocols Used for Inter-AS MVPN



25403

The following assumptions are made:

- PE-1 is the sender PE because the multicast source is directly connected to this router.
- PE-4 is the receiver PE because the multicast receiver is directly connected to this router.
- P-2 and P-3 are ASBR routers according to the Inter-AS model.

The multicast receiver and source can be indirectly connected to PE routers via CE routers, but for the core multicast distribution, these variations are conceptually the same. For simplicity, the PE and P router configurations are provided.

There are several challenges which have to be solved to make the complete inter-AS solution operational:

#### Challenge 1:

In case of Inter-AS MVPN Option B, routing information toward the source PE is not available in a remote AS domain because IGP routes are not exchanged between ASs.

As a result, a PIM-P join would never be sent upstream (from the receiver PE to the sender PE in a different AS). However, the PIM-P join has to be propagated from PE-4 to PE-1. Therefore, a solution is required to issue PIM-P join and perform RPF.

#### Solution:

Use a PIM reverse path forwarding (RPF) vector (RPFV) to propagate PIM-P over multiple segments. In this example there are three segments:

- PE-4 -> ASBR P-3
- ASBR P-3 -> ASBR P-2
- ASBR P-2 -> PE-1

The RPF vector is added to a PIM join by the PE router when the following option is enabled:

```
# on PE-4
configure router pim rpfv ?
- no rpfv [core] [mvpn]
- rpfv core mvpn
- rpfv core
```



```
- rpfv mvpn
<core>          : Proxy RPF vector for core
<mvpn>         : Proxy RPF vector for inter-AS rosen mvpn
```

The **mvpn** keyword enables "mvpn RPF vector" processing for Inter-AS Option B MVPN based on RFC 5496 and RFC 6513. If a core RPF vector is received, it is dropped before a message is processed.

All routers on the multicast traffic transport path must have this option enabled to allow RPF vector processing. If the option is not enabled, the RPF vector is dropped and the PIM join is processed as if the RPF vector is not present.

Details about RPF vector can be found in the following RFCs: 5496, 5384, 6513.

### Challenge 2:

With Inter-AS MVPN Option B, the BGP next-hop is modified by the local and remote ASBRs during re-advertisement of VPN IPv4 routes. When the BGP next-hop is changed, information about the originator of the prefix is lost when the advertisement reaches the receiver PE node. Therefore, a solution is required to do a successful RPF check for the VPN source at receiver VPRN.

This challenge does not apply to Model C because in Model C the BGP next-hop for VPN routes is not updated.

### Solution:

The transitive BGP connector attribute is added and used to advertise an address of a sender PE node which is carried inside a VPN IPv4 update. The BGP connector attribute allows the sender PE address information to be available to the receiver PE so that a receiver PE is able to associate VPN IPv4 advertisement to the corresponding source PE.

Inter-AS Option B works when the following criteria are met:

- Rosen MVPN is used with PIM SSM
- BGP MDT-SAFI address family is used
- PIM RPF vector is configured
- BGP connector attribute is used for VPN-IPv4 updates

SR OS inter-AS Option B is designed to be standard compliant based on the following RFCs:

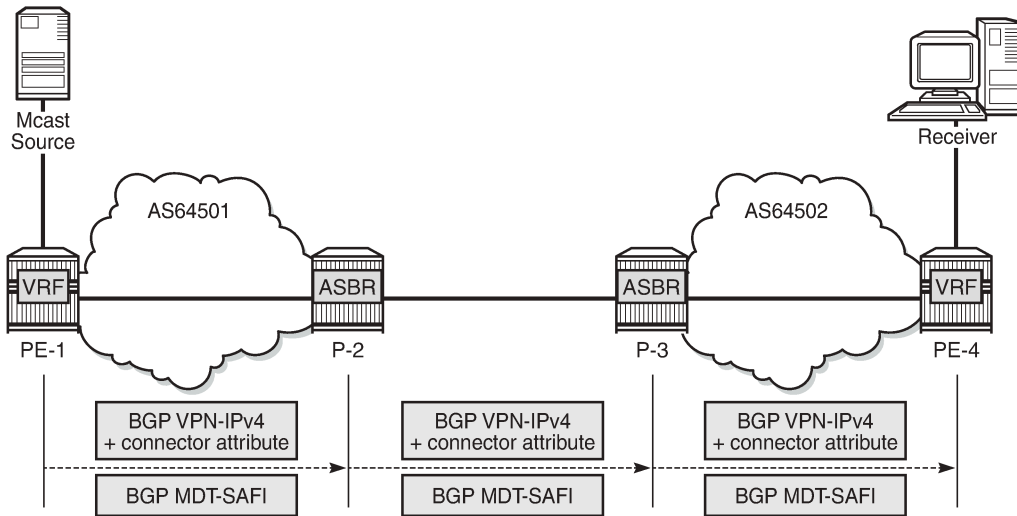
- RFC 5384, *The Protocol Independent Multicast (PIM) Join Attribute Format*
- RFC 5496, *The Reverse Path Forwarding (RPF) Vector TLV*
- RFC 6513, *Multicast in MPLS/BGP IP VPNs*

The following signaling stages can be identified when Inter-AS MVPN is configured:

- Stage 1 - BGP core signaling
- Stage 2 - Core PIM signaling
- Stage 3 - Customer PIM signalling

**Stage 1** - BGP core signaling

Figure 70: BGP Signaling Steps



25404

The sender PE sends VPN-IPv4 and MDT-SAFI BGP updates for this particular MVPN:

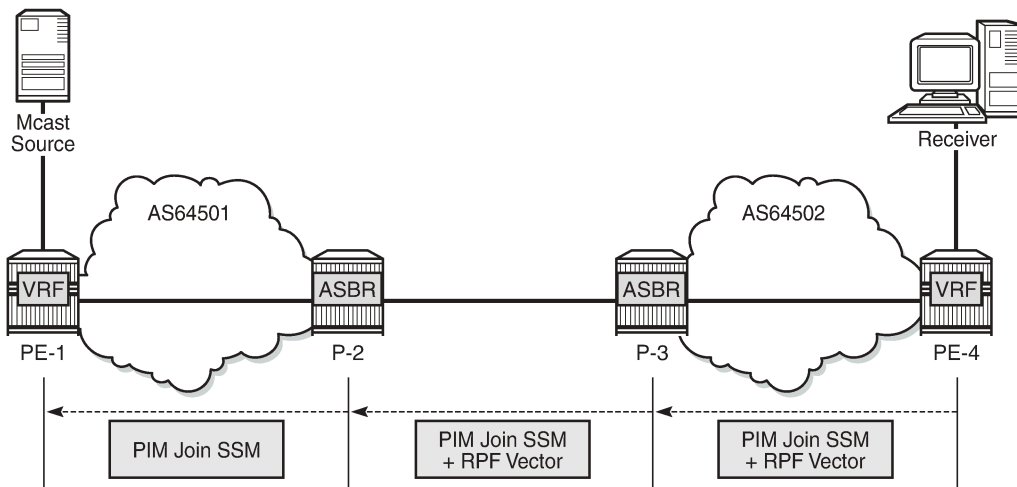
- Every ASBR propagates VPN-IPv4 and MDT-SAFI BGP updates:
  - Next hop (NH) attribute is modified every time
  - Connector attribute stays untouched

When this stage is completed, all routers have information necessary:

- to start PIM signaling in the core network (PIM-P) to prepare the default MDT
- to start PIM signaling of customer multicast streams (PIM-C) inside the VPN

**Stage 2 - Core PIM signaling**

Figure 71: PIM-P Signaling Steps for Default MDT



25405

PE-4 determines the reverse path to the source based on the RPF vector (ASBR P-3 IP address) and not based on the IP address of the multicast source (PE-1) which is unknown to PE-4.

PE-4 inserts an RPF vector and sends a PIM-P join to the immediate next-hop to reach ASBR P-3. Intermediate P-routers (if present) do not change the RPF vector.

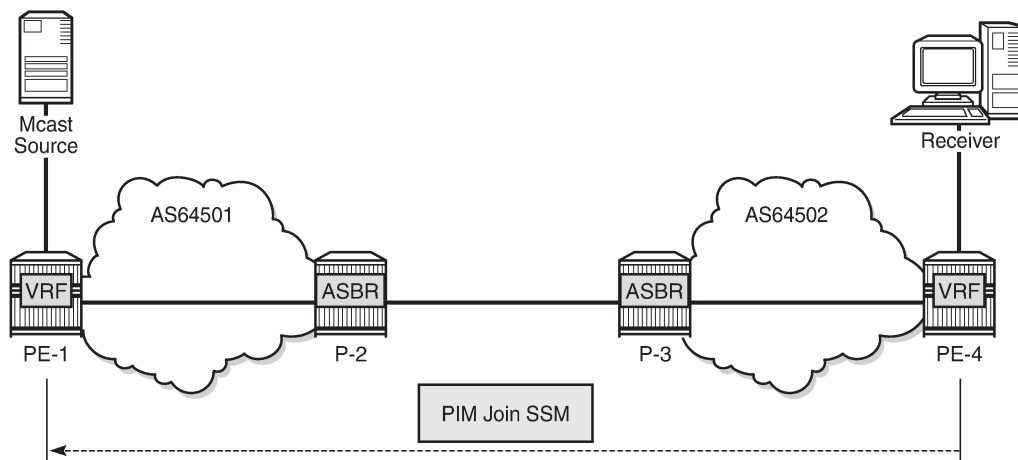
P-3 finds itself in the RPF vector and has to make a decision based on MDT-SAFI BGP table:

- P-3 determines the reverse path to the multicast source based on the RPF vector (ASBR P-2 IP address).
- If the multicast source and the NH do not match, P-3 has to use the RPFV.
- P-3 modifies the PIM-P join received from PE-4 with ASBR P-2's IP address as the upstream (taken from next hop MDT-SAFI network layer reachability information (NLRI)).
- P-2 can match the source IP with the NH in BGP MDT-SAFI. Therefore, there is no need for the RPF vector to be used.
- P-2 removes the RPF vector and sends a normal PIM-P join toward PE-1.

When this stage is completed, the default MDT is established for this MVPN and PE routers have the necessary information to start PIM signaling inside the VPRN (PIM-C).

### Stage 3 - Customer PIM signaling

Figure 72: PIM-C Signaling



25406

A PIM-C join is sent to the source PE using the existing tunnel infrastructure to the RPF neighbor PE-1 provided by the BGP connector attribute of the vpn-ipv4 route of the multicast source.

When this stage is completed, the customer multicast flows throughout the network in a default MDT.

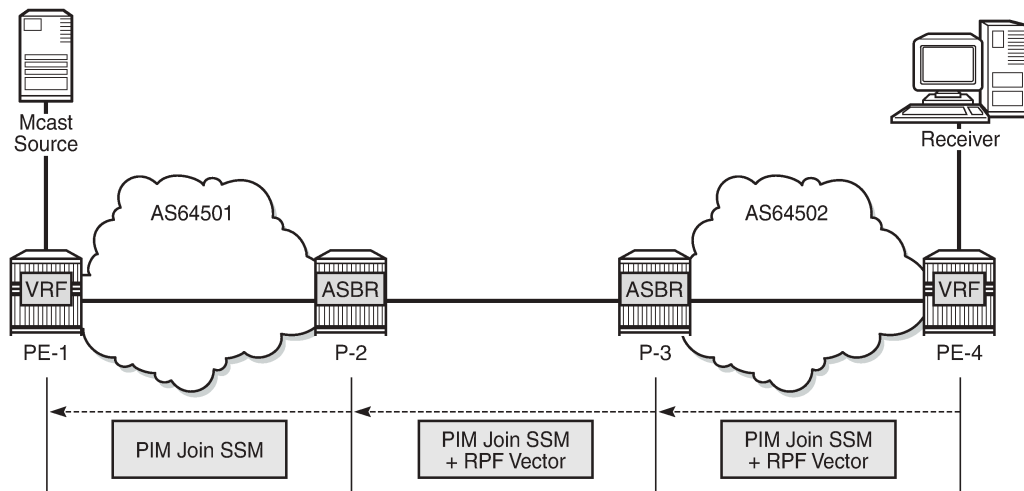
### Stage 4 - The multicast stream threshold is reached.

This stage is optional and applicable when S-PMSI instance and S-PMSI threshold are configured.

The process is similar to the default MDT setup:

- PE-4 determines the reverse path to the source based on the RPF vector (ASBR P-3's IP address) and not based on the IP address of the multicast source (PE-1) which is unknown to PE-4.
- PE-4 inserts an RPF vector and sends a PIM-P Join to the immediate next hop to reach ASBR P-3.

Figure 73: PIM-P Signaling Steps for Data MDT



25405

- Intermediate P-routers (if present) do not change the RPF vector.
- P-3 finds itself in the RPF vector and has to make a decision based on the MDT-SAFI BGP table:
  - P-3 determines the reverse path to the multicast source based on the RPF Vector (ASBR P-2's IP address).
  - If the multicast source and the NH do not match, P-3 has to use the RPFV.
  - P-3 modifies the PIM-P join received from PE-4 with ASBR P-2's IP address as upstream (taken from next hop MDT-SAFI NLRI).
- P-2 can match the source IP with the NH in the BGP MDT-SAFI. Therefore, there is no need for the RPF vector to be used.
- P-2 removes the RPF vector and sends a normal PIM-P join toward PE-1.

When this optional stage is completed, the customer multicast traffic flows through a dedicated Data MDT.

The SR OS implementation was also designed to interoperate with Cisco routers' Inter-AS implementations that do not fully comply with the RFC 5384 and RFC 5496.

When the following option is enabled:

```
configure router pim rpfv mvpn
```

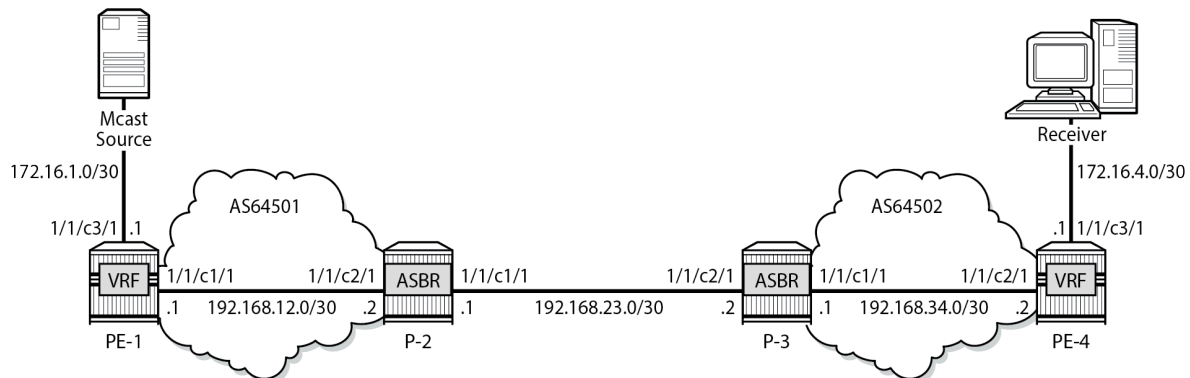
Cisco routers need to be configured to include **rd** in an RPF vector using the following command for interoperability:

```
ip multicast vrf <name> rpf proxy rd vector
```

## Configuration

The example topology is shown in [Figure 74: Example Topology Details](#).

Figure 74: Example Topology Details



25408b

The following components are used in the example scenario:

- VPRN 1
- Customer multicast group is 232.0.0.0/8
- Default MDT multicast group is 239.255.0.1
- Data MDT multicast group is 239.255.1.0/24
- Multicast source is 172.16.1.2
- PE-x routers have system IP addresses 192.0.2.x
- P-x routers have system IP addresses 192.0.2.x
- Interface between Router A and B has IP address 192.168.AB.x

Global BGP configuration for PE-1 router using the mdt-safi family with an iBGP neighbor to P-2. The system interface IP address is used for the iBGP session.

```
# on PE-1
configure
router
  bgp
    group "iBGP"
      family vpn-ipv4 mdt-safi
      type internal
      neighbor 192.0.2.2
      next-hop-self
    exit
  exit
exit
```

The global BGP configuration for P-2 router is using the mdt-safi family with an iBGP neighbor to PE-1 and an eBGP neighbor to P-3. The system interface IP address is used for the iBGP session and the network interface IP address is used for the eBGP session.

```
# on P-2
configure
router
  bgp
```

```

enable-inter-as-vpn
group "eBGP"
  family vpn-ipv4 mdt-safi
  neighbor 192.168.23.2
    type external
    peer-as 64502
  exit
exit
group "iBGP"
  family vpn-ipv4 mdt-safi
  neighbor 192.0.2.1
    next-hop-self
    type internal
  exit
exit
exit

```

The global BGP configuration for the router P-3 is using the mdt-safi family with an iBGP neighbor to PE-4 and an eBGP neighbor to P-2. The system interface IP address is used for the iBGP session and the network interface IP address is used for the eBGP session.

```

# on P-3
configure
router
  bgp
    enable-inter-as-vpn
    group "eBGP"
      family vpn-ipv4 mdt-safi
      neighbor 192.168.23.1
        type external
        peer-as 64501
      exit
    exit
    group "iBGP"
      family vpn-ipv4 mdt-safi
      neighbor 192.0.2.4
        next-hop-self
        type internal
      exit
    exit
  exit

```

The global BGP configuration for router PE-4 is using the mdt-safi family with an iBGP neighbor to P-3. The system interface IP address is used for the iBGP session.

```

# on PE-4
configure
router
  bgp
    group "iBGP"
      family vpn-ipv4 mdt-safi
      type internal
      neighbor 192.0.2.3
        next-hop-self
      exit
    exit
  exit

```

The global PIM configuration for all routers is as follows:

```

# on all routers

```

```
configure
router
pim
  rpf-table both
  apply-to non-ies
  rpfv mvpn
exit
```

The VPRN configuration for the PE routers is as follows:

```
# on PE-1
configure
service
  vprn 1 name "1" customer 1 create
  interface "int-PE-1-S-1" create
  address 172.16.1.1/30
  sap 1/1/c3/1 create
  exit
exit
  bgp-ipvpn
  mpls
    auto-bind-tunnel
    resolution-filter
    ldp
    rsvp
    exit
    resolution filter
  exit
  route-distinguisher 1:1
  vrf-target target:1:1
  no shutdown
  exit
exit
  pim
  apply-to all
  no shutdown
  exit
  mvpn
  auto-discovery mdt-safi
  provider-tunnel
  inclusive
  pim ssm 239.255.0.1
  exit
  exit
  selective
  data-threshold 232.0.0.0/8 1
  pim-ssm 239.255.1.0/24
  exit
  exit
  vrf-target unicast
  exit
  exit
  no shutdown
  exit
exit
```

```
# on PE-4
configure
service
  vprn 1 name "1" customer 1 create
  interface "int-PE-4-H-4" create
  address 172.16.4.1/30
  sap 1/1/c3/1 create
```

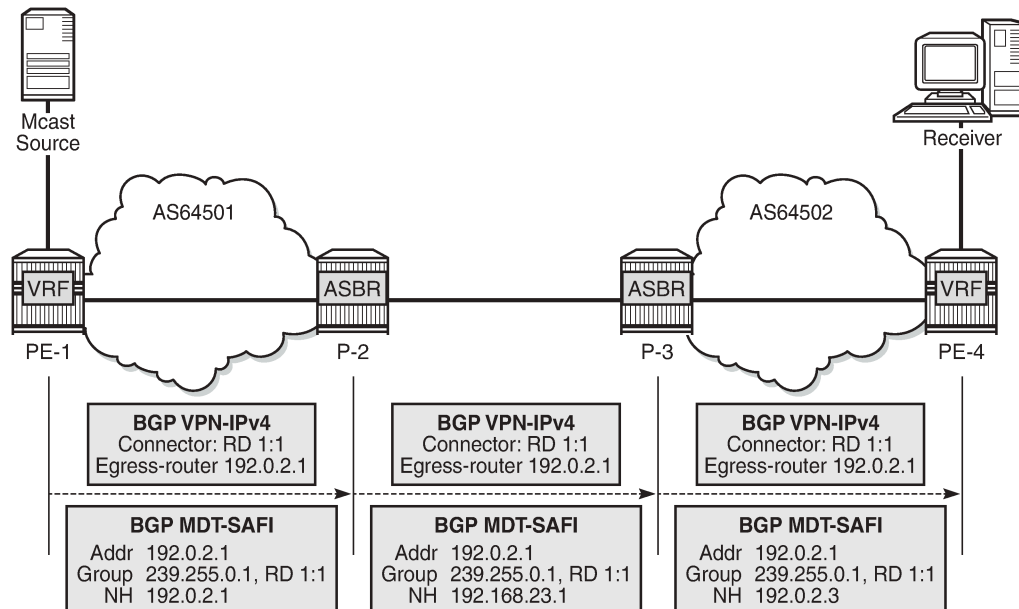
```
    exit
  exit
  bgp-ipvpn
  mpls
    auto-bind-tunnel
    resolution-filter
    ldp
    rsvp
    exit
    resolution filter
  exit
  route-distinguisher 4:1
  vrf-target target:1:1
  no shutdown
  exit
exit
igmp
  interface "int-PE-4-H-4"
  exit
exit
pim
  apply-to all
  no shutdown
exit
mvpn
  auto-discovery mdt-safi
  provider-tunnel
    inclusive
      pim ssm 239.255.0.1
    exit
  selective
    data-threshold 232.0.0.0/8 1
    pim-ssm 239.255.1.0/24
  exit
  exit
  vrf-target unicast
  exit
exit
no shutdown
exit
exit
```



## MVPN Verification and Debugging

### BGP Core Signaling

Figure 75: BGP Signaling Steps



25409

On PE-1, the **debug router bgp update** output shows the BGP update messages which are sent to P-2. The VPN-IPv4 update contains a connector attribute and the MDT-SAFI update is used for signaling multicast group 239.255.0.1.

```

1 2023/11/15 22:26:21.394 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 = 79
  Flag: 0x90 Type: 14 Len: 33 Multiprotocol Reachable NLRI:
    Address Family VPN_IPV4
    NextHop len 12 NextHop 192.0.2.1
    172.16.1.0/30 RD 1:1 Label 524285 (Raw label 0x7fffd1)
  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: 8 Extended Community:
    target:1:1
  Flag: 0xc0 Type: 20 Len: 14 Connector:
    RD 1:1, Egress-router 192.0.2.1
"

2 2023/11/15 22:26:21.394 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 = 62
Flag: 0x90 Type: 14 Len: 26 Multiprotocol Reachable NLRI:
  Address Family MDT-SAFI
    NextHop len 4 NextHop 192.0.2.1
    [MDT-SAFI] Addr 192.0.2.1, Group 239.255.0.1, RD 1:1
Flag: 0x40 Type: 1 Len: 1 Origin: 0
Flag: 0x40 Type: 2 Len: 0 AS Path:
Flag: 0x80 Type: 4 Len: 4 MED: 0
Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
Flag: 0xc0 Type: 16 Len: 8 Extended Community:
  target:1:1
"

```

On P-2, the **debug router bgp update** output shows the BGP update messages which are sent to P-3. The VPN-IPv4 update contains an unmodified connector attribute and the MDT-SAFI update is used for signaling multicast group 239.255.0.1.

```

3 2023/11/15 22:26:37.914 UTC MINOR: DEBUG #2001 Base Peer 1: 192.168.23.2
"Peer 1: 192.168.23.2: UPDATE
Peer 1: 192.168.23.2 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 78
  Flag: 0x90 Type: 14 Len: 33 Multiprotocol Reachable NLRI:
    Address Family VPN_IPV4
      NextHop len 12 NextHop 192.168.23.1
      172.16.1.0/30 RD 1:1 Label 524285 (Raw label 0x7fffd1)
Flag: 0x40 Type: 1 Len: 1 Origin: 0
Flag: 0x40 Type: 2 Len: 6 AS Path:
  Type: 2 Len: 1 < 64501 >
Flag: 0xc0 Type: 16 Len: 8 Extended Community:
  target:1:1
  Flag: 0xc0 Type: 20 Len: 14 Connector:
    RD 1:1, Egress-router 192.0.2.1
"

4 2023/11/15 22:26:37.914 UTC MINOR: DEBUG #2001 Base Peer 1: 192.168.23.2
"Peer 1: 192.168.23.2: UPDATE
Peer 1: 192.168.23.2 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 54
  Flag: 0x90 Type: 14 Len: 26 Multiprotocol Reachable NLRI:
    Address Family MDT-SAFI
      NextHop len 4 NextHop 192.168.23.1
      [MDT-SAFI] Addr 192.0.2.1, Group 239.255.0.1, RD 1:1
Flag: 0x40 Type: 1 Len: 1 Origin: 0
Flag: 0x40 Type: 2 Len: 6 AS Path:
  Type: 2 Len: 1 < 64501 >
Flag: 0xc0 Type: 16 Len: 8 Extended Community:
  target:1:1
"

```

On P-3, the **debug router bgp update** output shows the BGP update messages which are sent to PE-4. The VPN-IPv4 update contains an unmodified connector attribute and the MDT-SAFI update is used for signaling multicast group 239.255.0.1.

```

7 2023/11/15 22:26:52.726 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.4
"Peer 1: 192.0.2.4: UPDATE
Peer 1: 192.0.2.4 - Send BGP UPDATE:
  Withdrawn Length = 0
  Total Path Attr Length = 85
  Flag: 0x90 Type: 14 Len: 33 Multiprotocol Reachable NLRI:
    Address Family VPN_IPV4

```

```

NextHop len 12 NextHop 192.0.2.3
172.16.1.0/30 RD 1:1 Label 524285 (Raw label 0x7fffd1)
Flag: 0x40 Type: 1 Len: 1 Origin: 0
Flag: 0x40 Type: 2 Len: 6 AS Path:
Type: 2 Len: 1 < 64501 >
Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
Flag: 0xc0 Type: 16 Len: 8 Extended Community:
target:1:1
Flag: 0xc0 Type: 20 Len: 14 Connector:
RD 1:1, Egress-router 192.0.2.1
"

8 2023/11/15 22:26:52.726 UTC MINOR: DEBUG #2001 Base Peer 1: 192.0.2.4
"Peer 1: 192.0.2.4: UPDATE
Peer 1: 192.0.2.4 - Send BGP UPDATE:
Withdrawn Length = 0
Total Path Attr Length = 61
Flag: 0x90 Type: 14 Len: 26 Multiprotocol Reachable NLRI:
Address Family MDT-SAFI
NextHop len 4 NextHop 192.0.2.3
[MDT-SAFI] Addr 192.0.2.1, Group 239.255.0.1, RD 1:1
Flag: 0x40 Type: 1 Len: 1 Origin: 0
Flag: 0x40 Type: 2 Len: 6 AS Path:
Type: 2 Len: 1 < 64501 >
Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
Flag: 0xc0 Type: 16 Len: 8 Extended Community:
target:1:1
"

```

The BGP tables on PE-1 and PE-4 are updated accordingly. The most interesting aspect here is the MDT-SAFI routes received.

PE-4 has one MDT-SAFI update received from PE-1. The next-hop was modified according to the Option-B model.

```

*A:PE-4# show router bgp neighbor 192.0.2.3 received-routes mdt-safi
=====
BGP Router ID:192.0.2.4      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP MDT-SAFI Routes
=====
Flag Network                               LocalPref  MED
  Nexthop                               Group-Addr  IGP Cost
  As-Path                                Label
-----
u*>i 1:1:192.0.2.1                          100        None
      192.0.2.3                            239.255.0.1
64501
-----
Routes : 1
=====

```

PE-1 has one MDT-SAFI update received from PE-4. The next-hop was modified according to the Option B model.

```

*A:PE-1# show router bgp neighbor 192.0.2.2 received-routes mdt-safi

```

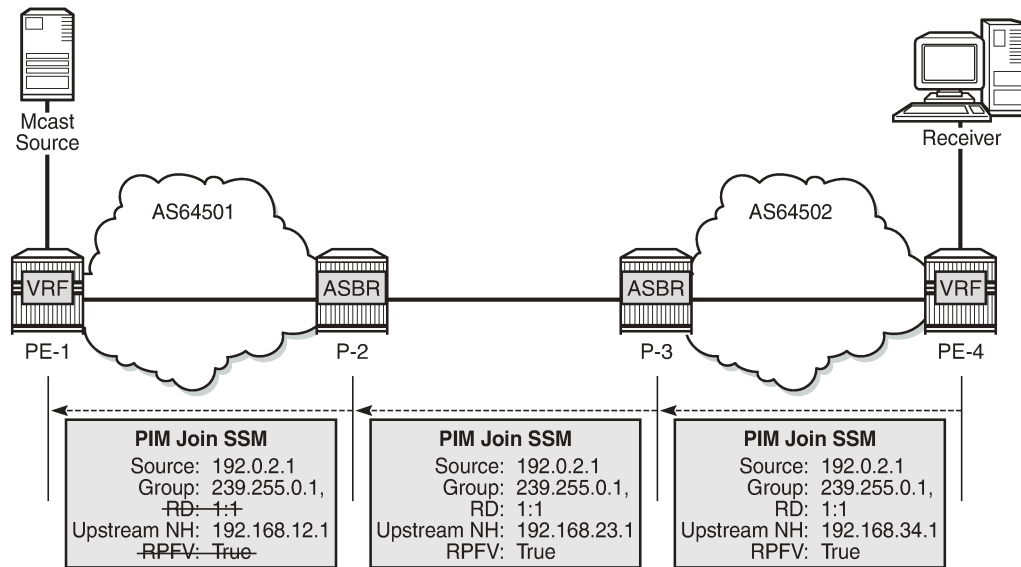
```

=====
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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP MDT-SAFI Routes
=====
Flag  Network                LocalPref  MED
Nexthop                Group-Addr  IGP Cost
As-Path                Label
-----
u*>i  4:1:192.0.2.4          100        None
      192.0.2.2            239.255.0.1
64502
-----
Routes : 1
=====

```

## Core PIM Signaling

Figure 76: PIM-P Signaling Steps for Default MDT



25410

On PE-4, the **debug router pim packet jp** output shows the PIM join/prune message which is sent to P-3. This message contains the original source of the multicast traffic (PE-1: 192.0.2.1) and the RPF Vector (P-3: 192.0.2.3).

```

8 2023/11/15 22:29:34.986 UTC MINOR: DEBUG #2001 Base PIM[Instance 1 Base]
"PIM[Instance 1 Base]: Join/Prune
[000 08:02:57.070] PIM-TX ifId 2 ifName int-PE-4-P-3 0.0.0.0 -> 224.0.0.13 Length: 48
PIM Version: 2 Msg Type: Join/Prune Checksum: 0x3d2c
Upstream Nbr IP : 192.168.34.1 Resvd: 0x0, Num Groups 1, HoldTime 210

```

```
Group: 239.255.0.1/32 Num Joined Srcs: 1, Num Pruned Srcs: 0
Joined Srcs:
192.0.2.1/32 Flag S <S,G> JA={rpfvMvpn 192.0.2.3 1:1}
"
```

On P-3, the **debug router pim packet jp** output shows the PIM join/prune message which is propagated to P-2. The source of multicast traffic is untouched while the RPF Vector is modified for Inter-AS propagation.

```
17 2023/11/15 22:29:34.950 UTC MINOR: DEBUG #2001 Base PIM[Instance 1 Base]
"PIM[Instance 1 Base]: Join/Prune
[000 08:02:45.990] PIM-TX ifId 2 ifName int-P-3-P-2 0.0.0.0 -> 224.0.0.13 Length: 48
PIM Version: 2 Msg Type: Join/Prune Checksum: 0x3286
Upstream Nbr IP : 192.168.23.1 Resvd: 0x0, Num Groups 1, HoldTime 210
Group: 239.255.0.1/32 Num Joined Srcs: 1, Num Pruned Srcs: 0
Joined Srcs:
192.0.2.1/32 Flag S <S,G> JA={rpfvMvpn 192.168.23.1 1:1}
"
```

On P-2, the **debug router pim packet jp** output shows the PIM join/prune message which is propagated to PE-1. The source of the multicast traffic is untouched while the RPF Vector is not present anymore.

```
16 2023/11/15 22:29:35.559 UTC MINOR: DEBUG #2001 Base PIM[Instance 1 Base]
"PIM[Instance 1 Base]: Join/Prune
[000 08:02:44.680] PIM-TX ifId 2 ifName int-P-2-PE-1 0.0.0.0 -> 224.0.0.13 Length: 34
PIM Version: 2 Msg Type: Join/Prune Checksum: 0x563f
Upstream Nbr IP : 192.168.12.1 Resvd: 0x0, Num Groups 1, HoldTime 210
Group: 239.255.0.1/32 Num Joined Srcs: 1, Num Pruned Srcs: 0
Joined Srcs:
192.0.2.1/32 Flag S <S,G>
"
```

As a result of this signaling, the default MDT is established between the two ASs. This can be checked with the **show router pim group** command.

The following PE-1 output shows the active multicast groups which are used as default MDT.

```
*A:PE-1# show router pim group
=====
Legend:  A = Active   S = Standby
=====
PIM Groups ipv4
=====
Group Address          Type          Spt Bit   Inc Intf   No.0ifs
  Source Address      RP           State     Inc Intf(S)
-----
239.255.0.1            (S,G)        spt       system     2
  192.0.2.1
239.255.0.1            (S,G)        spt       int-PE-1-P-2  1
  192.0.2.4
-----
Groups : 2
=====
```

The following PE-4 output shows the active multicast groups which are used as default MDT:

```
*A:PE-4# show router pim group
=====
Legend:  A = Active   S = Standby
=====
PIM Groups ipv4
=====
Group Address          Type          Spt Bit  Inc Intf      No.0ifs
  Source Address       RP
-----
239.255.0.1            (S,G)        spt      int-PE-4-P-3  1
  192.0.2.1
239.255.0.1            (S,G)        spt      system        2
  192.0.2.4
-----
Groups : 2
=====
```

The detailed information about the PIM-P group shows that the default MDT is used to deliver traffic. Key parameters such as the incoming/outgoing interfaces and non-zero traffic counters allow this conclusion to be made.

PE-4 has the incoming interface "int-PE-4-P-3", and outgoing interface "system", as follows:

```
*A:PE-4# show router pim group detail
=====
PIM Source Group ipv4
=====
Group Address       : 239.255.0.1
Source Address      : 192.0.2.1
RP Address          : 0
Advt Router         : 192.0.2.3

Upstream RPFV Nbr   : 192.168.34.1
RPFV Type           : Mvpn 1:1          RPFV Proxy       : 192.0.2.3

Flags               : spt                Type              : (S,G)
Mode                : sparse
MRIB Next Hop       : 192.168.34.1
MRIB Src Flags      : remote
Keepalive Timer Exp: 0d 00:03:24
Up Time             : 0d 00:02:21      Resolved By       : rtable-u

Up JP State         : Joined              Up JP Expiry      : 0d 00:00:39
Up JP Rpt           : Not Joined StarG   Up JP Rpt Override : 0d 00:00:00

Register State      : No Info
Reg From Anycast RP: No

Rpf Neighbor        : 192.168.34.1
Incoming Intf      : int-PE-4-P-3
Outgoing Intf List : system

Curr Fwding Rate    : 0.000 kbps
Forwarded Packets   : 6                  Discarded Packets : 0
Forwarded Octets    : 468                 RPF Mismatches    : 0
Spt threshold       : 0 kbps              ECMP opt threshold : 7
Admin bandwidth     : 1 kbps

---snip---
```

```
-----  
Groups : 2  
=====
```

PE-1 has incoming the interface "system", and outgoing interfaces "system, int-PE-1-P-2", as follows:

```
*A:PE-1# show router pim group detail
```

```
=====
```

```
PIM Source Group ipv4
```

```
=====
```

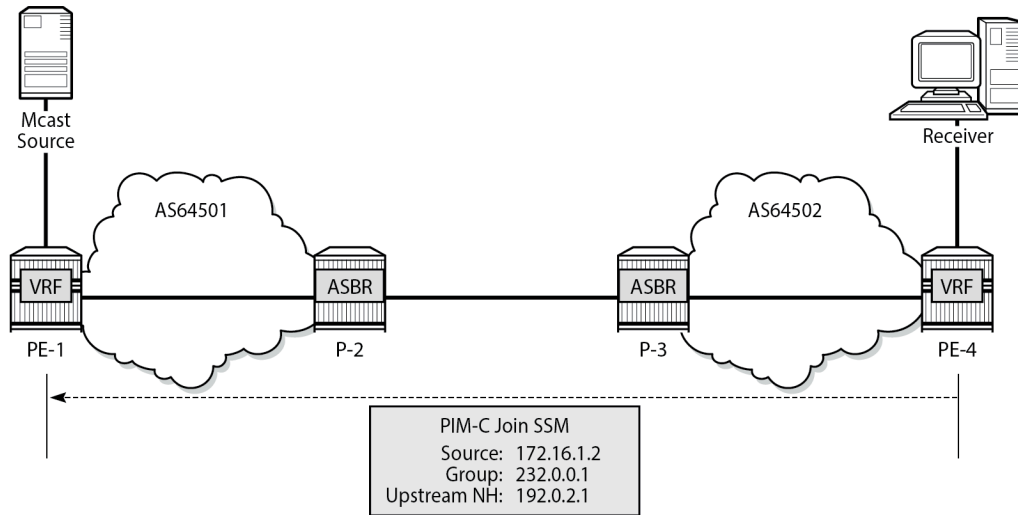
```
Group Address      : 239.255.0.1  
Source Address     : 192.0.2.1  
RP Address         : 0  
Advt Router        : 192.0.2.1  
Flags              : spt                Type           : (S,G)  
Mode               : sparse  
MRIB Next Hop      :  
MRIB Src Flags     : self  
Keepalive Timer Exp: 0d 00:03:08  
Up Time            : 0d 00:02:55      Resolved By      : rtable-m  
  
Up JP State        : Joined           Up JP Expiry      : 0d 00:00:04  
Up JP Rpt          : Not Joined StarG Up JP Rpt Override : 0d 00:00:00  
  
Register State     : No Info  
Reg From Anycast RP: No  
  
Rpf Neighbor       :  
Incoming Intf    : system  
Outgoing Intf List : system, int-PE-1-P-2  
  
Curr Fwding Rate   : 0.000 kbps  
Forwarded Packets  : 9                Discarded Packets : 0  
Forwarded Octets   : 702              RPF Mismatches    : 0  
Spt threshold      : 0 kbps           ECMP opt threshold : 7  
Admin bandwidth    : 1 kbps
```

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

```
-----  
Groups : 2  
=====
```

## Customer PIM Signaling

Figure 77: PIM-C Signaling



25411b

The PIM-C Join is sent to the sender PE using the existing tunnel infrastructure.

On PE-4, the **debug router 1 pim packet jp** output shows the PIM join/prune message which is sent to PE-1 using PMSI interface "1-mt-239.255.0.1" inside VPRN 1. All of this information and more can be found in the output of the **debug** command.

```
1 2023/11/15 22:34:30.579 UTC MINOR: DEBUG #2001 vprn1 PIM[Instance 2 vprn1]
"PIM[Instance 2 vprn1]: Join/Prune
[000 08:07:52.660] PIM-TX ifId 5798466 ifName 1-mt-239.255.0.1 0.0.0.0 -> 224.0.0.13 Length: 34
PIM Version: 2 Msg Type: Join/Prune Checksum: 0x7dd5
Upstream Nbr IP : 192.0.2.1 Resvd: 0x0, Num Groups 1, HoldTime 210
Group: 232.0.0.1/32 Num Joined Srcs: 1, Num Pruned Srcs: 0
Joined Srcs:
172.16.1.2/32 Flag S <S,G>
"
```

The detailed information about the PIM-C group for a particular VPRN shows that the default MDT is used to deliver traffic. For this purpose, the **show router 1 pim group detail** command is used. Key parameters such as the correct multicast group, correct incoming/outgoing interfaces and non-zero flow rate allow this conclusion to be made.

PE-1 has the incoming interface "int-PE-1-S-1", and outgoing interface "1-mt-239.255.0.1". If the threshold hasn't been reached to set up a selective provider tunnel, only one outgoing interface is listed. To generate this output, the data threshold for the selective provider tunnel was temporarily raised to 100000 kbps in VPRN 1.

```
*A:PE-1# show router 1 pim group detail
```

```
=====
PIM Source Group ipv4
=====
```



```

Group Address      : 232.0.0.1
Source Address    : 172.16.1.2
RP Address        : 0
Advt Router       : 192.0.2.1
Flags             :                               Type           : (S,G)
Mode              : sparse
MRIB Next Hop    : 172.16.1.2
MRIB Src Flags   : direct
Keepalive Timer   : Not Running
Up Time          : 0d 00:03:56      Resolved By         : rtable-u

Up JP State       : Joined           Up JP Expiry        : 0d 00:00:00
Up JP Rpt        : Not Joined StarG  Up JP Rpt Override  : 0d 00:00:00

Register State    : No Info
Reg From Anycast RP: No

Rpf Neighbor      : 172.16.1.2
Incoming Intf   : int-PE-1-S-1
Outgoing Intf List : 1-mt-239.255.0.1

Curr Fwding Rate  : 9325.072 kbps
Forwarded Packets : 280801           Discarded Packets   : 0
Forwarded Octets  : 275746582      RPF Mismatches      : 0
Spt threshold     : 0 kbps          ECMP opt threshold  : 7
Admin bandwidth   : 1 kbps
-----
Groups : 1
=====

```

PE-4 has the incoming interface "1-mt-239.255.0.1", and outgoing interface "int-PE-4-H-4" to the receiving host. As long as there is no S-PMSI, the following output can be seen.

```

*A:PE-4# show router 1 pim group detail

=====
PIM Source Group ipv4
=====
Group Address      : 232.0.0.1
Source Address    : 172.16.1.2
RP Address        : 0
Advt Router       : 192.0.2.3
Flags             :                               Type           : (S,G)
Mode              : sparse
MRIB Next Hop    : 192.0.2.1
MRIB Src Flags   : remote
Keepalive Timer   : Not Running
Up Time          : 0d 00:04:02      Resolved By         : rtable-u

Up JP State       : Joined           Up JP Expiry        : 0d 00:00:57
Up JP Rpt        : Not Joined StarG  Up JP Rpt Override  : 0d 00:00:00

Register State    : No Info
Reg From Anycast RP: No

Rpf Neighbor      : 192.0.2.1
Incoming Intf   : 1-mt-239.255.0.1
Outgoing Intf List : int-PE-4-H-4

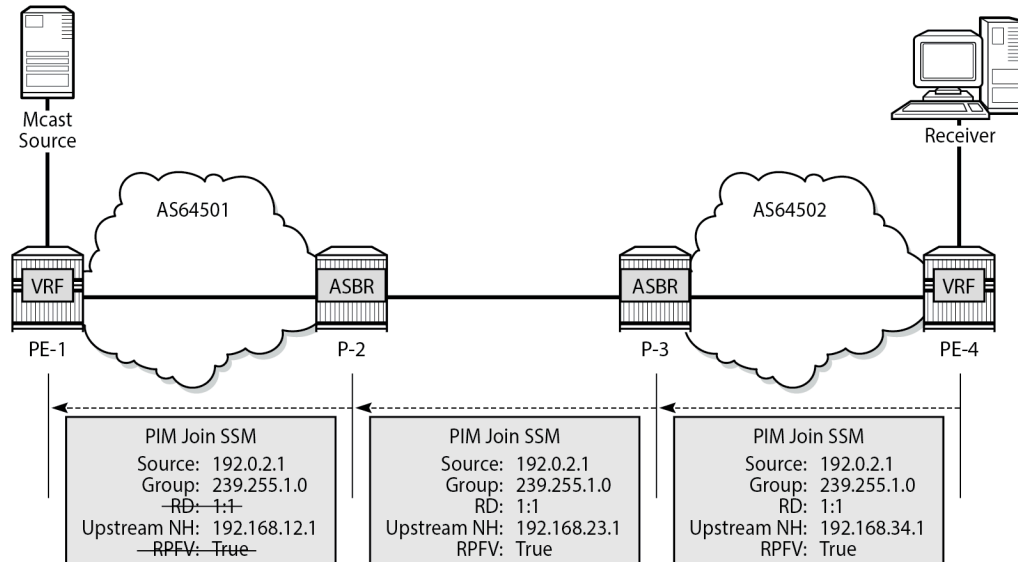
Curr Fwding Rate  : 9305.432 kbps
Forwarded Packets : 288185           Discarded Packets   : 0
Forwarded Octets  : 282997670      RPF Mismatches      : 0
Spt threshold     : 0 kbps          ECMP opt threshold  : 7

```

```
Admin bandwidth : 1 kbps
-----
Groups : 1
=====
```

## When Multicast Stream Threshold is Reached

Figure 78: PIM-P Signaling Steps for Data MDT



25412b

On PE-4, the **debug router pim packet jp** output shows the PIM join/prune message which is sent to P-3. This message contains the original source of the multicast traffic (PE-1: 192.0.2.1) and the RPF Vector (P-3: 192.0.2.3).

A new multicast group (239.255.1.0) is signaled for purposes of establishing the data MDT.

```
3 2023/11/15 22:40:51.643 UTC MINOR: DEBUG #2001 Base PIM[Instance 1 Base]
"PIM[Instance 1 Base]: Join/Prune
[000 08:14:13.730] PIM-TX ifId 2 ifName int-PE-4-P-3 0.0.0.0 -> 224.0.0.13 Length: 48
PIM Version: 2 Msg Type: Join/Prune Checksum: 0x3c2d
Upstream Nbr IP : 192.168.34.1 Resvd: 0x0, Num Groups 1, HoldTime 210
Group: 239.255.1.0/32 Num Joined Srcs: 1, Num Pruned Srcs: 0
Joined Srcs:
192.0.2.1/32 Flag S <S,G> JA={rpfvMvpn 192.0.2.3 1:1}
"
```

On P-3, the **debug router pim packet jp** output shows the PIM join/prune message which is propagated to P-2. The source of multicast traffic is untouched while the RPF Vector is modified for Inter-AS propagation.

```
6 2023/11/15 22:40:51.606 UTC MINOR: DEBUG #2001 Base PIM[Instance 1 Base]
"PIM[Instance 1 Base]: Join/Prune
[000 08:14:02.650] PIM-TX ifId 2 ifName int-P-3-P-2 0.0.0.0 -> 224.0.0.13 Length: 48
PIM Version: 2 Msg Type: Join/Prune Checksum: 0x3187
```

```
Upstream Nbr IP : 192.168.23.1 Resvd: 0x0, Num Groups 1, HoldTime 210
Group: 239.255.1.0/32 Num Joined Srcs: 1, Num Pruned Srcs: 0
Joined Srcs:
192.0.2.1/32 Flag S <S,G> JA={rpfvMvpn 192.168.23.1 1:1}
"
```

On P-2, the **debug router pim packet jp** output shows the PIM join/prune message which is propagated to PE-1. The source of multicast traffic is untouched while the RPF Vector is not present anymore.

```
6 2023/11/15 22:40:52.214 UTC MINOR: DEBUG #2001 Base PIM[Instance 1 Base]
"PIM[Instance 1 Base]: Join/Prune
[000 08:14:01.340] PIM-TX ifId 2 ifName int-P-2-PE-1 0.0.0.0 -> 224.0.0.13 Length: 34
PIM Version: 2 Msg Type: Join/Prune Checksum: 0x5540
Upstream Nbr IP : 192.168.12.1 Resvd: 0x0, Num Groups 1, HoldTime 210
Group: 239.255.1.0/32 Num Joined Srcs: 1, Num Pruned Srcs: 0
Joined Srcs:
192.0.2.1/32 Flag S <S,G>
"
```

As a result of this signaling, the Data MDT is established between the two ASs. This can be checked with **show router pim group** command.

The PE-1 output shows an additional multicast group (239.255.1.0), which was created in the global routing table (GRT).

```
*A:PE-1# show router pim group
=====
Legend:  A = Active   S = Standby
=====
PIM Groups ipv4
=====
Group Address          Type          Spt Bit   Inc Intf      No.0ifs
  Source Address       RP
-----
239.255.0.1            (S,G)         spt       system        2
  192.0.2.1
239.255.0.1            (S,G)         spt       int-PE-1-P-2  1
  192.0.2.4
239.255.1.0           (S,G)         system    1
  192.0.2.1
-----
Groups : 3
=====
```

The PE-4 output shows an additional multicast group (239.255.1.0), which was created in the GRT.

```
*A:PE-4# show router pim group
=====
Legend:  A = Active   S = Standby
=====
PIM Groups ipv4
=====
Group Address          Type          Spt Bit   Inc Intf      No.0ifs
  Source Address       RP
-----
239.255.0.1            (S,G)         spt       int-PE-4-P-3  1
  192.0.2.1
-----
```

```

239.255.0.1          (S,G)          spt      system    2
 192.0.2.4
239.255.1.0          (S,G)          int-PE-4-P-3  1
 192.0.2.1
-----
Groups : 3
=====

```

The detailed information about the PIM group in a VPRN shows that the data MDT is used to receive traffic instead of the default MDT.

The PE-4 output for multicast groups in a VPRN 1 has slightly changed: a new line "Incoming SPMSI Intf" was added. This indicates that the S-PMSI instance and dedicated Data MDT are used for this particular multicast group. The non-zero rate for the multicast flow is also an indication that multicast traffic is forwarded.

```

*A:PE-4# show router 1 pim group detail
=====
PIM Source Group ipv4
=====
Group Address       : 232.0.0.1
Source Address      : 172.16.1.2
RP Address          : 0
Advt Router        : 192.0.2.3
Flags               :                               Type           : (S,G)
Mode                : sparse
MRIB Next Hop      : 192.0.2.1
MRIB Src Flags     : remote
Keepalive Timer    : Not Running
Up Time            : 0d 00:02:15          Resolved By         : rtable-u

Up JP State        : Joined              Up JP Expiry        : 0d 00:00:44
Up JP Rpt         : Not Joined StarG    Up JP Rpt Override  : 0d 00:00:00

Register State     : No Info
Reg From Anycast RP: No

Rpf Neighbor       : 192.0.2.1
Incoming Intf      : 1-mt-239.255.0.1
Incoming SPMSI Intf: 1-mt-239.255.0.1*
Outgoing Intf List: int-PE-4-H-4

Curr Fwding Rate   : 9340.784 kbps
Forwarded Packets  : 161563              Discarded Packets   : 0
Forwarded Octets   : 158654866          RPF Mismatches      : 0
Spt threshold      : 0 kbps              ECMP opt threshold  : 7
Admin bandwidth    : 1 kbps

-----
Groups : 1
=====

```

The **show router 1 pim s-pmsi detail** command can also be used to verify existence of the S-PMSI instance for the VPRN 1. The output includes the multicast group inside the VPRN, the multicast source IP, the multicast group which is used for S-PMSI tunneling and the current forwarding rate.

```

*A:PE-4# show router 1 pim s-pmsi detail
=====
PIM Selective provider tunnels
=====

```

```
Md Source Address : 192.0.2.1      Md Group Address  : 239.255.1.0
Number of VPN SGs : 1              Uptime           : 0d 00:02:12
MT IfIndex        : 5806657        Egress Fwding Rate : 9340.784 kbps
Multistream-Id    : 0

VPN Group Address : 232.0.0.1
VPN Source Address : 172.16.1.2
State             : RX Joined
Expiry Timer      : 0d 00:01:55
=====
PIM Selective provider tunnels Interfaces : 1
=====
```

## Conclusion

Inter-AS MVPN offers flexibility for the operators who can use it to provide additional value added services to their customers. Before implementing this feature in the network the following are required:

- The RPF vector must be enabled on every router for inter-AS MVPN.
- Can be used only with Rosen MVPN with PIM SSM and MDT SAFI.

# Selective VPRN uRPF Control on Network Interfaces

This chapter provides information about selective VPRN uRPF control on network interfaces.

Topics in this chapter include:

- [Applicability](#)
- [Overview](#)
- [Configuration](#)
- [Conclusion](#)

## Applicability

The information and configuration in this chapter are based on SR OS Release 15.0.R7. Selective VPRN uRPF control on network interfaces is supported in SR OS Release 15.0.R1, and later.

## Overview

Unicast Reverse Path Forwarding (uRPF) can be used to reduce the vulnerability of networks to traffic flows with spoofed source IP addresses. By default, uRPF checking is disabled. In SR OS, uRPF can be enabled in loose mode or in strict mode on the ingress direction of both access and network interfaces.

- uRPF loose mode checking performs a longest-prefix match Forwarding Information Base (FIB) lookup on the IP source address (SA) of every incoming packet. All packets for which there is no matching non-default route are discarded.
- uRPF strict mode checking verifies, in addition to the check executed in loose mode, that the incoming interface matches the next-hop back toward the IP SA. Packets that enter on a different interface are discarded.



### Note:

Note: For VPRN traffic that is tunneled between routers, the route in the VPRN FIB has no interface associated with the prefix, so it is impossible to determine whether a packet with a spoofed source IP address enters the router via the correct interface. In that case, uRPF checking is identical for both modes: spoofed source IP address packets with no matching non-default route in the VPRN FIB are discarded.

This chapter focuses on network interfaces. The following command is used to enable uRPF checking on a network interface for IPv4 traffic.

```
*A:PE-1# configure router interface "int-PE-1-PE-2" urpf-check
```

The following command is used to enable uRPF checking on a network interface for IPv6 traffic:

```
*A:PE-1# configure router interface "int-PE-1-PE-2" ipv6 urpf-check
```

The default uRPF mode is strict. The uRPF mode can be changed as follows.

```
*A:PE-1# configure router interface "int-PE-1-PE-2" urpf-check mode ?
- mode {strict|loose|strict-no-ecmp}
```

When enabled on a base router network interface, uRPF operates as follows.

- For packets arriving on the network interface that require forwarding in the base router, uRPF checking performs a lookup of the IP SA in the base router FIB.
- For packets arriving on the network interface that require forwarding in a VPRN, uRPF checking performs a lookup of the IP SA in the VPRN FIB for locally configured VPRNs.

In some cases, uRPF checking should not be performed for all locally configured VPRNs, for example for VPRNs with asymmetric routing, such as when PE-1 has a route toward PE-2, but PE-2 has no route back to PE-1. Selective VPRN uRPF control on network interfaces offers the possibility to define for which locally configured VPRNs the uRPF should be checked. The following two commands control this selective or per-VPRN uRPF approach:

1. The first command is the following network interface-level command:

```
*A:PE-1# configure router interface "int-PE-1-PE-2" urpf-selected-vprns
```

2. The second command is the following VPRN-specific command that indicates this VPRN should be included in the set of VPRNs covered by the preceding **urpf-selected-vprns** command.

```
*A:PE-1# configure service vprn 1 network ingress urpf-check
```

When a specific VPRN should be excluded from the selective VPRN uRPF check, **no urpf-check** must be configured explicitly within that **vprn** context. Excluding a VPRN from uRPF checking only works for the network interfaces with **urpf-selected-vprns** enabled and **urpf-check mode value** configured. When uRPF is configured on a network interface without **urpf-selected-vprns**, uRPF checking is inherited by all locally configured VPRNs, regardless of the presence of the **configure vprn <service-id> network ingress urpf-check** command.

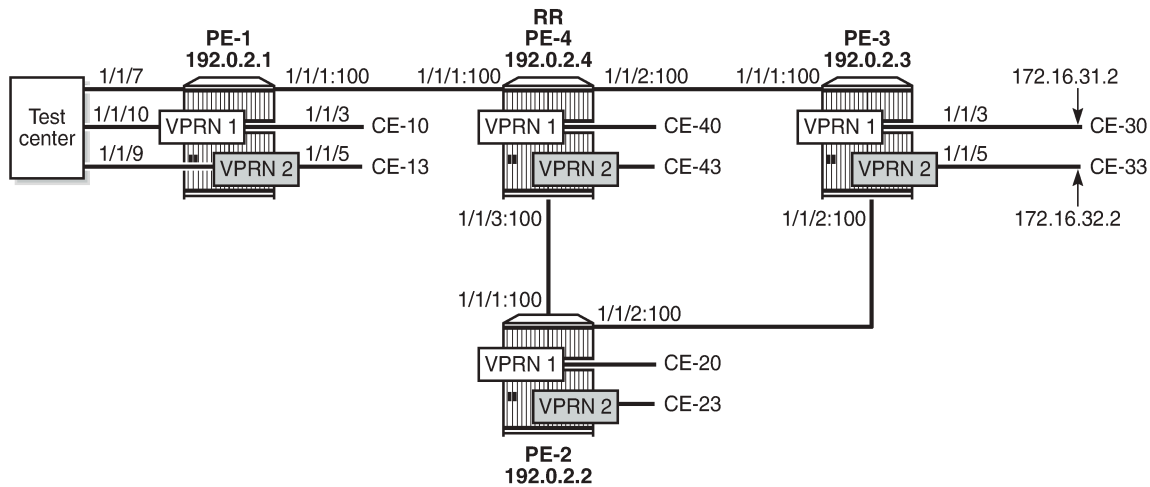
The uRPF checking behavior is as follows.

- If uRPF is enabled on the network interface without **urpf-selected-vprns**, the existing uRPF behavior applies.
- If uRPF is disabled on the network interface, but **urpf-selected-vprns** is enabled, no uRPF lookup is done for any packet arriving on the network interface.
- If uRPF is enabled on the network interface and **urpf-selected-vprns** is enabled, uRPF checking is performed on all packets to be forwarded in the base router. For packets to be forwarded in a VPRN, uRPF checking is only performed for locally configured VPRNs that have **urpf-check** enabled on the network ingress in the VPRN.

## Configuration

[Figure 79: Example Topology in AS 64496](#) shows the example topology with four routers and a test center. On each of the routers, VPRN 1 and VPRN 2 are configured. The test center, connected to PE-1, can send IPv4 or IPv6 traffic toward the base router (port 1/1/7) and both VPRNs (port 1/1/10 for VPRN 1 and port 1/1/9 for VPRN 2).

Figure 79: Example Topology in AS 64496



27503

The initial configuration on the four PEs includes the following:

- Cards, MDAs
- Ports:
  - Network ports between the PEs and access ports toward the VPRNs on each PE
  - Port 1/1/7 on PE-1 toward the test center is configured as a network port and is associated with a network interface in the base router. Ports 1/1/9 and 1/1/10 on PE-1 toward the test center are configured as access ports and associated with SAPs in the VPRNs.
- Network interfaces between the PEs and from PE-1 to the test center with a dual-stack IPv4/IPv6
- IS-IS as IGP (alternatively, OSPF can be used) on all network interfaces
- LDP on all network interfaces between the PEs

As an example, the initial configuration on PE-1 is as follows. By default, uRPF is disabled on the network interfaces. The configuration on the other PEs is similar.

```
*A:PE-1# configure
router
  interface "int-PE-1-PE-4"
    address 192.168.14.1/30
    port 1/1/1:100
    ipv6
      address 2001:db8::14:1/126
    exit
  exit
  interface "int-PE-1-TestCenter"
    address 192.168.11.1/30
    port 1/1/7
    ipv6
      address 2001:db8::11:1/126
    exit
  exit
  interface "system"
    address 192.0.2.1/32
    ipv6
```



```

        address 2001:db8::2:1/128
    exit
exit
isis
    area-id 49.0001
    ipv6-routing native
    interface "system"
    exit
    interface "int-PE-1-PE-4"
        interface-type point-to-point
    exit
    no shutdown
exit
ldp
    interface-parameters
        interface "int-PE-1-PE-4" dual-stack
            ipv4
                no shutdown
            exit
        exit
    exit
exit
exit

```

The initial VPRN service configuration on PE-1 is as follows. Auto-bind-tunnel is enabled and LDP tunnels will be used. The service configuration on the other PEs is similar, excluding the interface to the test center.

```

configure
  service
    vprn 1 customer 1 create
      description "PE-1-VPRN-1"
      route-distinguisher 64496:1
      auto-bind-tunnel
        resolution any
      exit
      vrf-target target:64496:1
      interface "int-PE-1-CE-10" create
        address 172.16.11.1/24
        ipv6
          address 2001:db8::11:1/120
        exit
        sap 1/1/3 create
        exit
      exit
      interface "toTestCenter" create
        address 172.16.110.1/24
        ipv6
          address 2001:db8::110:1/120
        exit
        sap 1/1/10 create
        exit
      exit
      no shutdown
    exit
    vprn 2 customer 1 create
      description "PE-1-VPRN-2"
      route-distinguisher 64496:2
      auto-bind-tunnel
        resolution any
      exit
      vrf-target target:64496:2
      interface "int-PE-1-CE-13" create
        address 172.16.12.1/24
        ipv6

```

```

        address 2001:db8::12:1/120
        exit
        sap 1/1/5 create
        exit
    exit
    interface "toTestCenter" create
        address 172.16.120.1/24
        ipv6
            address 2001:db8::120:1/120
        exit
        sap 1/1/9 create
        exit
    exit
    no shutdown
exit

```

BGP is configured for the VPN-IPv4 and VPN-IPv6 address families with PE-4 as route reflector. The following is the BGP configuration on PE-1:

```

configure
  router
    autonomous-system 64496
    bgp
      split-horizon
      group "iBGP"
        family vpn-ipv4 vpn-ipv6
        peer-as 64496
        neighbor 192.0.2.4
      exit
    exit
  exit
exit

```

In this example, no uRPF checking will be enabled on the access interfaces of the VPRNs, but obviously, it might be combined with uRPF control on network interfaces.

The following commands to enable uRPF in strict mode (default) on the access interface are only shown for completeness:

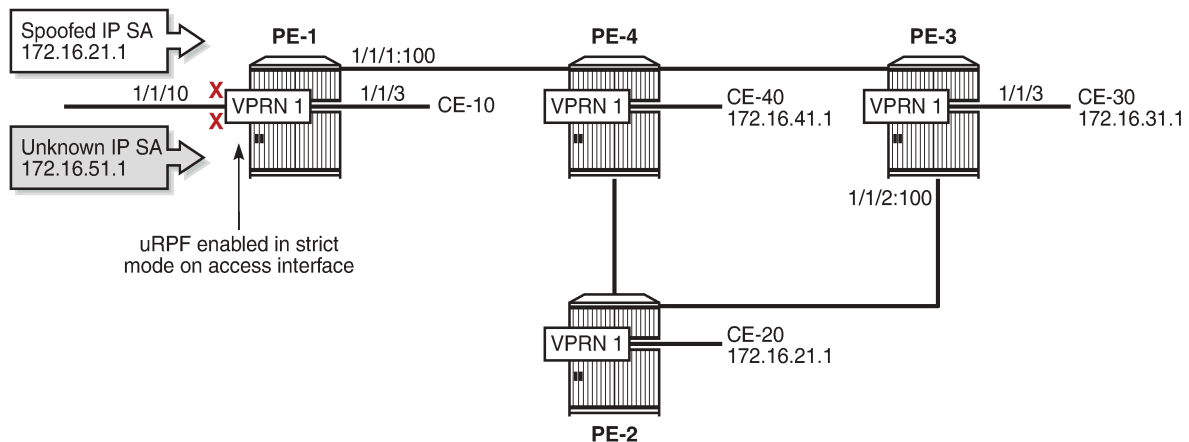
```

*A:PE-1# configure service vprn 1 interface "toTestCenter" urpf-check
*A:PE-1# configure service vprn 1 interface "toTestCenter" ipv6 urpf-check

```

With this configuration, packets with spoofed or unknown IP SAs arriving on a VPRN SAP, where uRPF checking is enabled in strict mode, are dropped at the access interface, as shown in [Figure 80: uRPF Enabled in Strict Mode on Access Interface in VPRN 1](#). IP SA 172.16.21.1 has a non-default route in the FIB via a tunnel to PE-2, but packets with this IP SA are not expected on the access interface, so they are dropped in strict mode on interface "toTestCenter" (whereas they would be forwarded in loose mode). All packets with unknown IP SAs—for which there is no non-default route in the FIB of the VPRN—are dropped in strict and in loose mode on interface "toTestCenter".

Figure 80: uRPF Enabled in Strict Mode on Access Interface in VPRN 1



27504

In the remainder of this chapter, uRPF is disabled on the access interfaces, so all packets with spoofed or unknown IP SAs in VPRN 1 will be forwarded by PE-1.

uRPF is enabled on the network interfaces of all PEs. The following commands enable uRPF (in strict mode, by default) for IPv4 and IPv6 on the network interfaces on PE-1. The configuration is similar on the other PEs.

```
*A:PE-1# configure router interface "int-PE-1-PE-4" urpf-check
*A:PE-1# configure router interface "int-PE-1-PE-4" ipv6 urpf-check
*A:PE-1# configure router interface "int-PE-1-TestCenter" urpf-check
*A:PE-1# configure router interface "int-PE-1-TestCenter" ipv6 urpf-check
```

The FIB for the base router on PE-1 is as follows.

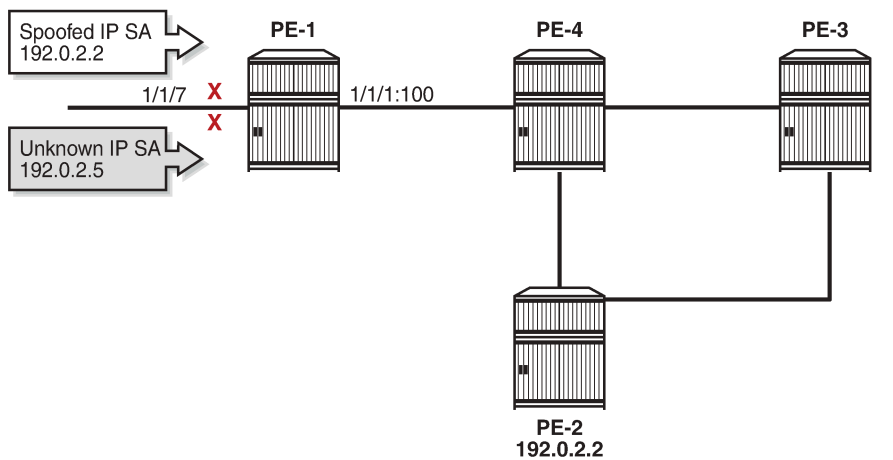
```
*A:PE-1# show router fib 1

=====
FIB Display
=====
Prefix [Flags]                               Protocol
NextHop
-----
192.0.2.1/32                                  LOCAL
  192.0.2.1 (system)
192.0.2.2/32                                  ISIS
  192.168.14.2 (int-PE-1-PE-4)
192.0.2.3/32                                  ISIS
  192.168.14.2 (int-PE-1-PE-4)
192.0.2.4/32                                  ISIS
  192.168.14.2 (int-PE-1-PE-4)
192.168.11.0/30                               LOCAL
  192.168.11.0 (int-PE-1-TestCenter)
192.168.14.0/30                               LOCAL
  192.168.14.0 (int-PE-1-PE-4)
192.168.23.0/30                               ISIS
  192.168.14.2 (int-PE-1-PE-4)
192.168.24.0/30                               ISIS
  192.168.14.2 (int-PE-1-PE-4)
192.168.34.0/30                               ISIS
  192.168.14.2 (int-PE-1-PE-4)
```

-----  
Total Entries : 9

The test center sends two traffic flows with IP destination address (DA) 192.0.2.3 to the base router on PE-1. The first traffic flow has IP SA 192.0.2.2, which is the system address of PE-2 and is expected on another network interface, so it will be dropped by uRPF in strict mode. The second traffic flow has IP SA 192.0.2.5 for which there is no non-default route available in the FIB, so it will be dropped due to uRPF checking. [Figure 81: uRPF Checking in Strict Mode in Base Router on PE-1](#) shows how uRPF drops packets with spoofed or unknown IP SAs at the incoming network interface "int-PE-1-TestCenter" on PE-1.

Figure 81: uRPF Checking in Strict Mode in Base Router on PE-1



27505

The following monitor command output on PE-1 shows that the incoming traffic on network port 1/1/7 toward the test center is dropped. No traffic is forwarded to port 1/1/1 toward PE-4. The packets sent and received on port 1/1/1 are of a different nature, such as IS-IS messages.

```
*A:PE-1# monitor port 1/1/1 1/1/7 rate interval 3 repeat 2
=====
Monitor statistics for Ports
=====
-----snip-----
At time t = 3 sec (Mode: Rate)
-----
Port 1/1/1
-----
Octets                214                140
Packets                2                   2
Errors                 0                   0
Bits                  1712               1120
Utilization (% of port capacity)  ~0.00             ~0.00

Port 1/1/7
-----
Octets                270251              0
Packets               2111                0
Errors                 0                   0
Bits                  2162008             0
```

```
Utilization (% of port capacity)          2.49          0.00
-----
---snip---
```

The IPv6 FIB on PE-1 is as follows.

```
*A: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::628:1ff:fe01:1 (int-PE-1-PE-4)
2001:db8::2:3/128                             ISIS
  fe80::628:1ff:fe01:1 (int-PE-1-PE-4)
2001:db8::2:4/128                             ISIS
  fe80::628:1ff:fe01:1 (int-PE-1-PE-4)
2001:db8::11:0/126                            LOCAL
  2001:db8::11:0 (int-PE-1-TestCenter)
2001:db8::14:0/126                            LOCAL
  2001:db8::14:0 (int-PE-1-PE-4)
2001:db8::23:0/126                            ISIS
  fe80::628:1ff:fe01:1 (int-PE-1-PE-4)
2001:db8::24:0/126                            ISIS
  fe80::628:1ff:fe01:1 (int-PE-1-PE-4)
2001:db8::34:0/126                            ISIS
  fe80::628:1ff:fe01:1 (int-PE-1-PE-4)
-----
Total Entries : 9
```

Similar results occur for IPv6 traffic with IP DA 2001:db8::2:3 and IP SA 2001:db8::2:2 (system IPv6 address of PE-2) or IP SA 2001:db8::2:5 (unknown IP SA). The following port statistics show that the packets are dropped at the incoming port 1/1/7 toward the test center instead of being forwarded to port 1/1/1 toward PE-4. Instead of using the port statistics, the preceding monitor command can also be used.

```
*A:PE-1# clear port 1/1/[1..10] statistics
*A:PE-1# sleep 2
*A:PE-1# show port 1/1/[1..10] statistics

=====
Port Statistics on Slot 1
=====
Port      Ingress      Ingress      Egress      Egress
Id        Packets      Octets       Packets     Octets
-----
1/1/1          3           426          2           253
=====

=====
Port Statistics on Slot 1
=====
Port      Ingress      Ingress      Egress      Egress
Id        Packets      Octets       Packets     Octets
-----
1/1/7      4236         542208       0            0
=====
```

```
*A:PE-1#
```

## uRPF Control on Network Interfaces Inherited by VPRNs

By default, the uRPF control settings of the network interface are inherited by the VPRNs.

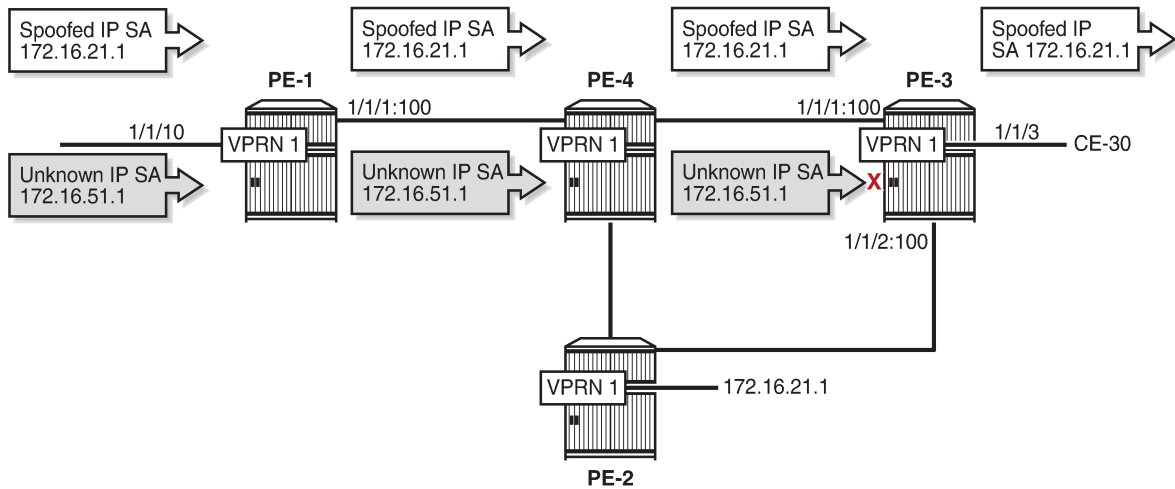
The test center sends a first traffic flow with IP DA 172.16.31.2 (CE-30) to SAP 1/1/10 of VPRN 1 on PE-1. The traffic flow has IP SA 172.16.21.1, which has a non-default route in the FIB of VPRN 1 on all PEs. Afterward, the test center sends a second traffic flow with IP DA 172.16.31.2 (CE-30) to SAP 1/1/10 of VPRN 1 on PE-1. These packets have IP SA 172.16.51.1, which is unknown in the VPRN FIB. uRPF is disabled on the access interface, so the packets are not dropped at the SAP, but forwarded in tunnels toward PE-3. No uRPF checking is performed on PE-4, because it is not the endpoint of the tunnel. The tunnel terminates at PE-3 and uRPF is checked on the incoming network interface. The FIB for VPRN 1 on PE-3 is as follows.

```
*A:PE-3# show router 1 fib 1
```

```
=====
FIB Display
=====
Prefix [Flags]                                Protocol
NextHop
-----
172.16.11.0/24                                BGP_VPN
  192.0.2.1 (VPRN Label:262135 Transport:LDP)
172.16.21.0/24                                BGP_VPN
  192.0.2.2 (VPRN Label:262135 Transport:LDP)
172.16.31.0/24                                LOCAL
  172.16.31.0 (int-CE-31-CE-30)
172.16.41.0/24                                BGP_VPN
  192.0.2.4 (VPRN Label:262135 Transport:LDP)
172.16.110.0/24                              BGP_VPN
  192.0.2.1 (VPRN Label:262135 Transport:LDP)
-----
Total Entries : 5
```

All IP packets with IP SA 172.16.21.1 will be forwarded to CE-30, regardless of the interface where they are received, because no network interface is associated with prefix 172.16.21.0/24 in the FIB of VPRN 1. [Figure 82: uRPF Checking in VPRN 1 on PE-3](#) shows that the only packets that will be dropped on PE-3 have an IP SA for which no non-default route is present in the FIB of VPRN 1; in this case, IP SA 172.16.51.1.

Figure 82: uRPF Checking in VPRN 1 on PE-3



27506

The following monitor command output on PE-3 for the traffic flow with IP DA 172.16.31.2 and IP SA 172.16.21.1 shows that the traffic is forwarded to port 1/1/3 toward CE-30.

```
*A:PE-3# monitor port 1/1/1 1/1/3 rate interval 3 repeat 2
=====
Monitor statistics for Ports
=====
Input                                     Output
-----
---snip---
At time t = 3 sec (Mode: Rate)
-----
Port 1/1/1
-----
Octets                                     295703                                   174
Packets                                   2112                                    1
---snip---
Port 1/1/3
-----
Octets                                     0                                       270251
Packets                                   0                                       2111
---snip---
```

The following monitor command output on PE-3 for the traffic flow with IP DA 172.16.31.2 and unknown IP SA 172.16.51.1 shows that the traffic is dropped at ingress port 1/1/1 instead of being forwarded to port 1/1/3 toward CE-30.

```
*A:PE-3# monitor port 1/1/1 1/1/3 rate interval 3 repeat 2
=====
Monitor statistics for Ports
=====
Input                                     Output
-----
---snip---
```

```

-----
At time t = 3 sec (Mode: Rate)
-----
Port 1/1/1
-----
Octets                295630                154
Packets              2112                1
---snip---

Port 1/1/3
-----
Octets                0                0
Packets              0                0
---snip---

```

Similar results occur for IPv6 traffic flows toward CE-30 with spoofed or unknown IP SAs, but they are not included here. The IPv6 FIB for VPRN 1 on PE-3 is as follows.

```

*A:PE-3# show router 1 fib 1 ipv6

=====
FIB Display
=====
Prefix [Flags]                Protocol
NextHop
-----
2001:db8::11:0/120            BGP_VPN
  192.0.2.1 (VPRN Label:262135 Transport:LDP)
2001:db8::21:0/120            BGP_VPN
  192.0.2.2 (VPRN Label:262135 Transport:LDP)
2001:db8::31:0/120            LOCAL
  2001:db8::31:0 (int-CE-31-CE-30)
2001:db8::41:0/120            BGP_VPN
  192.0.2.4 (VPRN Label:262135 Transport:LDP)
2001:db8::110:0/120           BGP_VPN
  192.0.2.1 (VPRN Label:262135 Transport:LDP)
-----
Total Entries : 5

```

To show selective uRPF for different VPRNs, uRPF checking is needed on the network interfaces for VPRN 1, but not for VPRN 2. To achieve this, additional configuration is required to exclude VPRN 2 from the uRPF check. The following configuration in VPRN 2 is required, but not sufficient to exclude VPRN 2 from the uRPF check.

```

*A:PE-3# configure service vprn 2 network ingress no urpf-check

```

This setting is ignored because no selective VPRN uRPF checking is enabled on the network-interfaces level and the behavior remains unchanged: the uRPF settings are inherited by VPRN 2, even though the configuration in VPRN 2 might be misleading. When the test center generates a traffic flow with IP DA 172.16.32.2 (CE-33) and unknown IP SA 172.16.52.1, the traffic is dropped by PE-3 after uRPF checking. The FIB for VPRN 2 on PE-3 is as follows.

```

*A:PE-3# show router 2 fib 1

=====
FIB Display
=====
Prefix [Flags]                Protocol
NextHop
-----

```



```

172.16.12.0/24                                     BGP_VPN
  192.0.2.1 (VPRN Label:262134 Transport:LDP)
172.16.22.0/24                                     BGP_VPN
  192.0.2.2 (VPRN Label:262134 Transport:LDP)
172.16.32.0/24                                     LOCAL
  172.16.32.0 (int-CE-32-CE-33)
172.16.42.0/24                                     BGP_VPN
  192.0.2.4 (VPRN Label:262134 Transport:LDP)
172.16.120.0/24                                    BGP_VPN
  192.0.2.1 (VPRN Label:262134 Transport:LDP)
-----
Total Entries : 5

```

The following monitor command output on PE-3 shows that the traffic that enters network port 1/1/1 with unknown IP address is dropped; no packets are forwarded to port 1/1/5 toward CE-33. This implies that uRPF control is still active for VPRN 2.

```

*A:PE-3# monitor port 1/1/1 1/1/5 rate interval 3 repeat 2

=====
Monitor statistics for Ports
=====
-----
Input                                     Output
-----
---snip---
-----
At time t = 3 sec (Mode: Rate)
-----
Port 1/1/1
-----
Octets                                     295869                                     227
Packets                                    2114                                       2
---snip---
Port 1/1/5
-----
Octets                                     0                                           0
Packets                                    0                                           0
---snip---

```

A similar result occurs for IPv6 traffic toward CE-33 with IP DA 2001:db8::32:2 and unknown IP SA 2001:db8::52:1. The IPv6 FIB for VPRN 2 on PE-3 is as follows.

```

*A:PE-3# show router 2 fib 1 ipv6

=====
FIB Display
=====
Prefix [Flags]                               Protocol
NextHop
-----
2001:db8::12:0/120                            BGP_VPN
  192.0.2.1 (VPRN Label:262134 Transport:LDP)
2001:db8::22:0/120                            BGP_VPN
  192.0.2.2 (VPRN Label:262134 Transport:LDP)
2001:db8::32:0/120                            LOCAL
  2001:db8::32:0 (int-CE-32-CE-33)
2001:db8::42:0/120                            BGP_VPN
  192.0.2.4 (VPRN Label:262134 Transport:LDP)
2001:db8::120:0/120                           BGP_VPN
  192.0.2.1 (VPRN Label:262134 Transport:LDP)

```

-----  
Total Entries : 5

## Selective VPRN uRPF Control on Network Interfaces

Selective VPRN uRPF control on network interfaces requires the following:

- uRPF configured on the network interfaces (by default disabled): **urpf-check**
- Selective VPRN uRPF control enabled on the network interfaces: **urpf-selected-vprns** (by default disabled)
- **[no] urpf-check** configured on the network ingress of the VPRNs (by default enabled)

In this example, uRPF is already configured on the network interfaces. The configuration on PE-3 is as follows.

```
*A:PE-3# configure router interface "int-PE-3-PE-2" urpf-check
*A:PE-3# configure router interface "int-PE-3-PE-2" ipv6 urpf-check
*A:PE-3# configure router interface "int-PE-3-PE-4" urpf-check
*A:PE-3# configure router interface "int-PE-3-PE-4" ipv6 urpf-check
```

Selective VPRN uRPF control needs to be enabled on all nodes. The configuration on PE-3 is as follows.

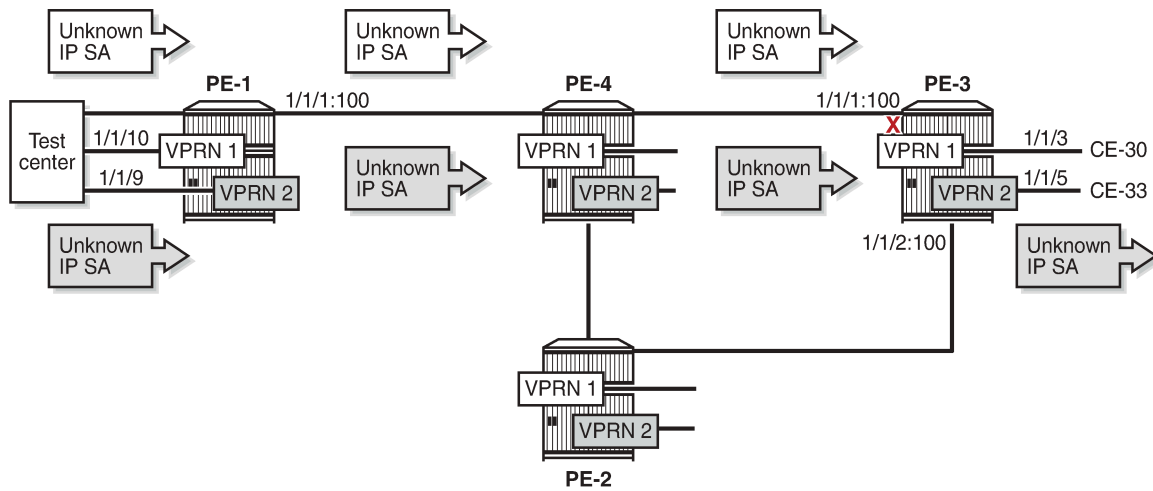
```
*A:PE-3# configure router interface "int-PE-3-PE-2" urpf-selected-vprns
*A:PE-3# configure router interface "int-PE-3-PE-4" urpf-selected-vprns
```

uRPF checking is enabled for VPRN 1 (default) and disabled for VPRN 2, as follows.

```
*A:PE-3# configure service vprn 2 network ingress no urpf-check
```

When the test center generates a traffic flow with IP DA 172.16.31.2 (CE-30) and unknown IP SA 172.16.51.1 in VPRN 1, the packets will be dropped at the incoming network port 1/1/1 on PE-3. For VPRN 2, traffic with IP DA 172.16.32.2 (CE-33) is forwarded, even if the IP SA is unknown (such as 172.16.52.1), because uRPF checking is disabled. [Figure 83: Selective VPRN uRPF on Network Interfaces Enabled for VPRN 1 and Disabled for VPRN 2](#) shows that packets with unknown IP SA in VPRN 1 are dropped by uRPF control on PE-3, while packets with unknown IP SA in VPRN 2 are forwarded on PE-3.

Figure 83: Selective VPRN uRPF on Network Interfaces Enabled for VPRN 1 and Disabled for VPRN 2



27507

The following monitor command output shows that traffic in VPRN 1 with IP DA 172.16.31.2 and IP SA 172.16.51.1 is dropped at incoming port 1/1/1 on PE-3. A similar result occurs for IPv6 addressing.

```
*A:PE-3# monitor port 1/1/1 1/1/3 rate interval 3 repeat 2
=====
Monitor statistics for Ports
=====
Input                               Output
-----
---snip---
At time t = 3 sec (Mode: Rate)
-----
Port 1/1/1
-----
Octets                               295667                               163
Packets                              2112                                1
---snip---
Port 1/1/3
-----
Octets                               0                                    0
Packets                              0                                    0
---snip---
```

The following monitor command output shows that traffic in VPRN 2 with IP DA 172.16.32.2 and IP SA 172.16.52.1 is forwarded to port 1/1/5 on PE-3 toward CE-33. A similar result occurs for IPv6 addressing.

```
*A:PE-3# monitor port 1/1/1 1/1/3 1/1/5 1/1/9 rate interval 3 repeat 2
=====
Monitor statistics for Ports
=====
Input                               Output
-----
---snip---
```

```
At time t = 3 sec (Mode: Rate)
-----
Port 1/1/1
-----
Octets                293565                186
Packets              2097                1
---snip---

Port 1/1/5
-----
Octets                0                268160
Packets              0                2095
---snip---
```

The uRPF control in the base router remains unchanged. In strict mode, PE-1 will drop all packets with spoofed or unknown IP addresses on the incoming network interface, as shown in [Figure 81: uRPF Checking in Strict Mode in Base Router on PE-1](#).

## Conclusion

uRPF checking can help service providers to mitigate spoofing attacks. uRPF checking can be executed for all base router traffic and VPRN traffic independently. When the routes held by specific VPRNs are asymmetric, it may be useful to exclude those VPRNs from network ingress uRPF checking.

# Spoke Termination for IPv6-6VPE

This chapter provides information about spoke termination for IPv6-6VPE.

Topics in this chapter include:

- [Applicability](#)
- [Overview](#)
- [Configuration](#)
- [Conclusion](#)

## Applicability

This chapter was originally written for SR OS Release 8.0, where Epipe Virtual Leased Line (VLL) is supported for IPv6 spoke termination within a Virtual Private Routed Network (VPRN). The CLI in the current edition corresponds to SR OS Release 21.10.R2.

## Overview

RFC 4659, *BGP-MPLS IP Virtual Private Network (VPN) Extension for IPv6 VPN*, standardized the use of an IPv6 over IPv4 tunneling scheme. SR OS supports the standardized IPv6 over IPv4 tunneling scheme for VPRN services using Multi-Protocol Border Gateway Protocol (MP-BGP), also known as 6VPE.

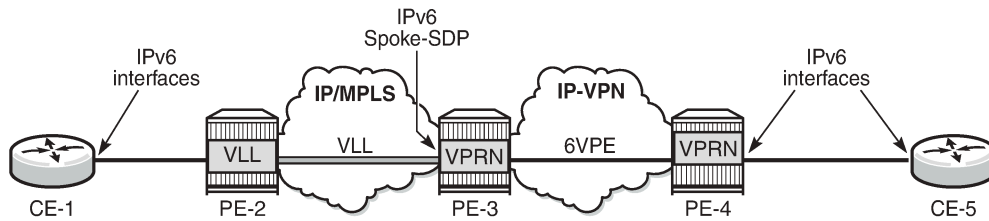
SR OS supports pseudowire termination by a VPRN from an Epipe Virtual Leased Line (VLL) or VPLS spoke service destination point (spoke SDP) where the pseudowire can be given IPv6 addresses and run IPv6 protocols. In the example used in this chapter, any advertisements across the Multi-Protocol Labeled Switching (MPLS) network between VPRN Provider Edge (PE) devices will use 6VPE. This chapter describes the configuration for IPv6 spoke termination to a VPRN over an Epipe VLL and transporting IPv6 packets over 6VPE tunnels between PE devices.

This solution can be used where a service provider is providing VPRN services built on a transport network whose Interior Gateway Protocol (IGP) is using IPv4 addressing on the network interfaces. The customer's CE and the service provider's PE must support IPv6 pseudowires, IPv6 interfaces and in addition, the service provider also must be able to support the advertisements of IPv6 prefixes between CE-PE peerings and between the transport PE routers using MP-BGP. The advertisement of IPv6 prefixes across the MPLS network and the transport of IPv6 traffic is tunneled using 6VPE.

The VPRN PE supports spoke termination of Epipe VLL services on access with IPv6 addressing between the CE and VPRN PE. The IPv6 spoke termination on VPRN services has the same functionality as VPRN IPv4 spoke termination.

The example in [Figure 84: Spoke termination for IPv6](#) illustrates a CE device that connects to a VPRN PE on an IPv6 interface addressing using spoke termination.

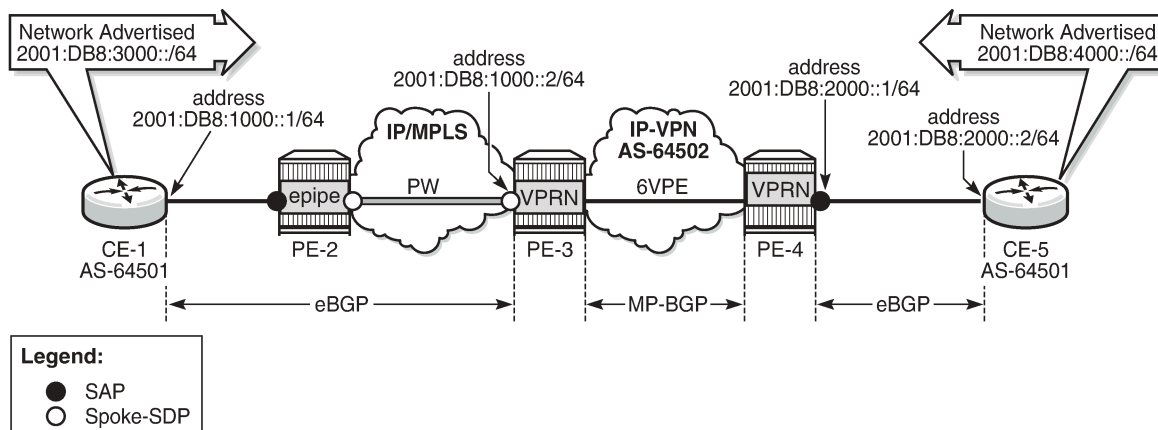
Figure 84: Spoke termination for IPv6



25455

CE-1 is connected to the VPRN service on PE-3, using IPv6 interfaces. CE-1 reaches PE-3 by connecting to PE-2. PE-2 uses an Epipe VLL for transport to the VPRN on the PE-3. The connectivity between the VLL service on the VPRN service on PE-3 is using spoke termination with IPv6 addressing on the spoke SDP interface on PE-3.

Figure 85: IPv6 addressing and IPv6 prefixes



25456

Figure 85: IPv6 addressing and IPv6 prefixes shows the overall IPv6 addressing from interfaces to prefixes advertised from CE-1 and CE-5 across the VPRN network.

- Link between CE-1 and PE-3: 2001:db8:1000::/64
- Link between CE-5 and PE-4: 2001:db8:2000::/64
- Advertised prefix from CE-1: 2001:db8:3000::/64
- Advertised prefix from CE-5: 2001:db8:4000::/64

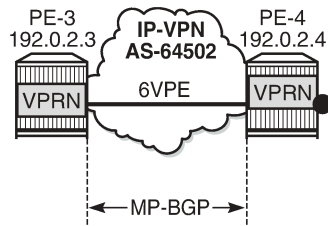
PE-3 has an MP-eBGP session with CE-1 to receive and advertise IPv6 routes. PE-3 also has an MP-iBGP peering session with PE-4 to use 6VPE to tunnel IPv6 routes and traffic to and from PE-4. PE-4 has an IPv6 SAP interface to CE-5 and uses MP-eBGP to advertise to and receive routes from CE-5 (no spoke termination). The configuration of PE-3 is included to provide examples of the end-to-end VPRN service using a 6VPE model.

This network topology illustrates the use of spoke termination using IPv6 interfaces and the tunneling of IPv6 traffic over a 6VPE MPLS network.

## Configuration

First an MPLS network is established where the VPRN service can use 6VPE to tunnel traffic across the IPv4 IGP.

Figure 86: MP-BGP VPN IPv6



25457

In [Figure 86: MP-BGP VPN IPv6](#), PE-3 and PE-4 are edge routers running VPRN services on access with IPv6 interfaces. The MPLS network is configured using IPv4 link addressing. Interior Border Gateway Protocol (iBGP) peerings need to be established with MP-BGP for the VPN-IPv6 address family between PE-3 and PE-4.

```
# on PE-3:
configure
router Base
  bgp
    group "iBGP"
      description "iBGP peering in AS 64502"
      family vpn-ipv6
      type internal
      neighbor 192.0.2.4
        description "PE-4"
      exit
    exit
  no shutdown
exit
exit
```

```
# on PE-4:
configure
router Base
  bgp
    group "iBGP"
      description "iBGP peering in AS 64502"
      family vpn-ipv6
      type internal
      neighbor 192.0.2.3
        description "PE-3"
      exit
    exit
  no shutdown
exit
exit
```

Configuring address family VPN-IPv6 between VPRN PE edge routers in BGP enables MP-BGP for the Layer 3 VPNs supporting the customer's IPv6 addressing (6VPE).

The following commands verify the BGP sessions for the VPN-IPv6 address family between PE-3 and PE-4:

```
*A:PE-3# show router bgp neighbor 192.0.2.4

=====
BGP Neighbor
=====
-----
Peer          : 192.0.2.4
Description   : PE-4
Group         : iBGP
-----
Peer AS       : 64502           Peer Port      : 50662
Peer Address  : 192.0.2.4
Local AS     : 64502           Local Port     : 179
Local Address : 192.0.2.3
Peer Type    : Internal       Dynamic Peer   : No
State        : Established    Last State     : Established
Last Event   : recvOpen
Last Error   : Cease (Connection Collision Resolution)
Local Family : VPN-IPv6
Remote Family: VPN-IPv6
---snip---
```

```
*A:PE-4# show router bgp neighbor 192.0.2.3

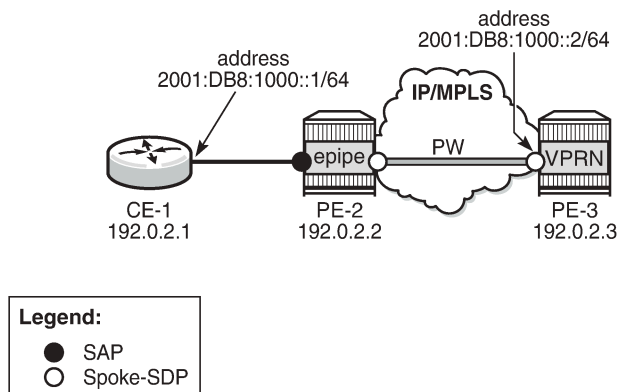
=====
BGP Neighbor
=====
-----
Peer          : 192.0.2.3
Description   : PE-3
Group         : iBGP
-----
Peer AS       : 64502           Peer Port      : 179
Peer Address  : 192.0.2.3
Local AS     : 64502           Local Port     : 50662
Local Address : 192.0.2.4
Peer Type    : Internal       Dynamic Peer   : No
State        : Established    Last State     : Active
Last Event   : recvOpen
Last Error   : Unrecognized Error
Local Family : VPN-IPv6
Remote Family: VPN-IPv6
---snip---
```

After the MP-BGP sessions are established for the VPN-IPv6 address-family, 6VPE tunnel support is provided between PE-3 and PE-4.

[Figure 87: Spoke termination for IPv6 addressing](#) illustrates the model for spoke termination for IPv6 using VPRN services.



Figure 87: Spoke termination for IPv6 addressing



25458

CE-1 is configured with IPv6 addressing on the access interface facing the VPRN service. CE-1's access is backhauled to the VPRN service on PE-3 using Epipe VLL with spoke termination. The configuration of the Epipe VLL on PE-2 is as follows:

```
# on PE-2:
configure
service
  sdp 231 mpls create
  far-end 192.0.2.3
  ldp
  no shutdown
exit
epipe 1 name "Epipe 1" customer 1 create
  sap 1/1/2 create
  no shutdown
  exit
  spoke-sdp 231:1 create
  no shutdown
  exit
  no shutdown
exit
exit
```

Epipe 1 on PE-2 is configured with a SAP interface facing the customer and a spoke SDP facing PE-3. The spoke SDP is terminated into the customer's VPRN service on PE-3.

The possible IPv6 options for spoke SDP interfaces on the CLI for VPRN Services are as follows (compliant with RFC 4213, *Basic Transition Mechanisms for IPv6 Hosts and Routers* <draft-ietf-v6ops-mech-v2-07.txt>):

- Interface spoke SDP (IPv6 options only)

```
*A:PE-3>config>service>vprn# interface "int-PE-3-PE-2" ?
---snip---
[no] ipv6          + Enables/Configures IPv6 for a VPRN interface
---snip---
```

```
*A:PE-3>config>service>vprn>if# ipv6 ?
- ipv6
- no ipv6
```

```
[no] address          - Assigns an IPv6 address to the VPRN interface.
[no] bfd              - Configure BFD parameters
[no] dad-disable      - Disable Duplicate Address Detection
[no] dhcp6-relay      + Configure DHCPv6 relay parameters for the VPRN interface
[no] dhcp6-server     + Configure DHCPv6 server parameters for the VPRN interface
[no] forward-ipv4-p*  - Enable/disable forwarding unencapsulated IPv4 packets
                    + Configure ICMPv6 parameters for the VPRN interface
[no] link-local-add*  - Configure link-local address
[no] local-dhcp-ser*  - Assign a DHCP server to the interface
[no] local-proxy-nd   - Enable/disable local proxy Neighbor Discovery on the VPRN
                    interface
                    + Configure ND host route to populate
[no] nd-learn-unsol*  - Configure neighbor discover learn unsolicited
[no] nd-proactive-r*  - Configure neighbor discovery proactive refresh
[no] neighbor         - Configure IPv6-to-MAC address mapping on the VPRN interface
[no] neighbor-limit   - Configures the maximum amount of IPv6 neighbor entries
[no] proxy-nd-policy  - Configure a proxy Neighbor Discovery policy for the VPRN
                    interface
[no] qos-route-look*  - Enable/Disable Qos route lookup for the interface
[no] reachable-time   - Configure neighbor reachability detection timer
[no] secure-nd        + Configure Secure Neighbor Discovery (SEND) parameters for the
                    interface
[no] stale-time       - Configure the time a neighbor discovery cache entry can remain
                    stale before being removed
[no] tcp-mss          - Configure TCP maximum segment size for the interface
[no] urpf-check       + Enables/Configures unicast RPF check for an interface
[no] vrrp             + Context to create and configure VRRP virtual router instance
                    on the interface
```

- IPv6 address

```
*A:PE-3>config>service>vprn>if>ipv6# address ?
- address <ipv6-address/prefix-length> [eui-64] [track-srrp <srrp-instance>]
  [modifier <cga-modifier>] [dad-disable][primary-preference <primary-preference>]
- no address <ipv6-address/prefix-length>

<ipv6-address/pref*> : ipv6-address  x:x:x:x:x:x:x:x  (eight 16-bit pieces)
                    x:x:x:x:x:x:d.d.d.d
                    x [0..FFFF]H
                    d [0..255]D
                    (no multicast address)
                    prefix-length [4..128]
<eui-64>             : keyword
<srrp-instance>     : [1..4294967295]
<cga-modifier>      : [0x0..0xFFFFFFFF... (32 hex nibbles)]
<dad-disable>       : keyword
<primary-preference> : [1..4294967295]
```

- DHCPv6 relay parameters for the VPRN service (default settings)

```
*A:PE-3>config>service>vprn>if>ipv6>dhcp6-relay# info detail
-----
shutdown
no description
no lease-populate
no neighbor-resolution
option
  no interface-id
  no remote-id
exit
no source-address
no link-address
```

```

no user-db
no python-policy
no server
-----

```

- DHCPv6 server parameters for the VPRN service (default)

```

*A:PE-3>config>service>vprn>if>ipv6>dhcp6-server# info detail
-----
prefix-delegation
shutdown
exit
max-nbr-of-leases 8000
-----

```

- ICMPv6 (default)

```

*A:PE-3>config>service>vprn>if>ipv6>icmp6# info detail
-----
packet-too-big 100 10
param-problem 100 10
redirects 100 10
time-exceeded 100 10
unreachables 100 10
-----

```

- Link-local-addressing, for the VPRN interface. By default, link-local addressing is assigned dynamically. Use this command if you want to add a static link-local-address.

```

*A:PE-3>config>service>vprn>if>ipv6# link-local-address ?
- link-local-address <ipv6-address> [dad-disable]
- no link-local-address

<ipv6-address>      : ipv6-address      - x:x:x:x:x:x:x
                                     x:x:x:x:x:d.d.d.d
                                     x [0..FFFF]H
                                     d [0..255]D

<dad-disable>      : keyword

```

- Neighbor: IPv6 to MAC address mapping on the VPRN interface

```

*A:PE-3>config>service>vprn>if>ipv6# neighbor ?
- neighbor <ipv6-address> <mac-address>
- no neighbor <ipv6-address>

<ipv6-address>      : x:x:x:x:x:x:x      (eight 16-bit pieces)
                                     x:x:x:x:x:d.d.d.d
                                     x [0..FFFF]H
                                     d [0..255]D
                                     prefix-length [1..128]
<mac-address>      : xx:xx:xx:xx:xx:xx or xx-xx-xx-xx-xx-xx

```

- Enabling local proxy neighbor discovery

```

*A:PE-3>config>service>vprn>if>ipv6# local-proxy-nd ?
- local-proxy-nd
- no local-proxy-nd

```

- VRRP

```
*A:PE-3>config>service>vprn>if>ipv6# vrrp ?
- no vrrp <virtual-router-id>
- vrrp <virtual-router-id> [owner] [passive]

<virtual-router-id> : [1..255]
<owner>              : keyword

[no] backup           - Configure virtual router IP addresses for the interface
[no] bfd-enable       - Configure a BFD interface
[no] init-delay       - Configure VRRP initialization delay timer
[no] mac              - Configure a Virtual MAC address to use in Neighbor Discovery
[no] master-int-inh*  - Allow/disallow the master instance to dictate the master down
                      timer (non-owner context only)
[no] message-interv*  - Configure the interval for sending VRRP Advertisement messages
[no] ntp-reply        - Allow/disallow non-owner master to reply to NTP requests
                      (non-owner context only)
[no] oper-group       - Associate group-name to VRRP
[no] ping-reply       - Allow/disallow non-owner master to reply to ICMP Echo requests
                      (non-owner context only)
[no] policy           - Associate a VRRP Priority Control Policy with the virtual
                      router instance (non-owner context only)
[no] preempt          - Allow/disallow the virtual router instance to override an
                      existing non-owner master (non-owner context only)
[no] priority         - Configure the base priority for the virtual router instance
                      (non-owner context only)
[no] shutdown         - Administratively enable/disable the virtual router instance
                      (non-owner context only)
[no] standby-forwar*  - Allow/disallow the forwarding of packets by a standby router
[no] telnet-reply     - Allow/disallow non-owner master to reply to Telnet requests
                      (non-owner context only)
[no] traceroute-rep*  - Allow/disallow non-owner master to reply to traceroute requests
                      (non-owner context only)
```

The VPRN on PE-3 exports IPv6 routes (IPv6 route on CE-5) to CE-1 using the following route policy.

```
# on PE-3
configure
router Base
  policy-options
  begin
  prefix-list "PE-3-CE-1"
    prefix 2001:db8:4000::/64 exact
  exit
  policy-statement "PE-3-BGP-CE-1"
    entry 10
      from
        prefix-list "PE-3-CE-1"
      exit
      action accept
        origin igp
      exit
    exit
  default-action drop
  exit
  exit
  commit
exit
```

The configuration for the VPRN service on PE-3 with IPv6 interface (spoke SDP) as shown in [Figure 87: Spoke termination for IPv6 addressing](#):

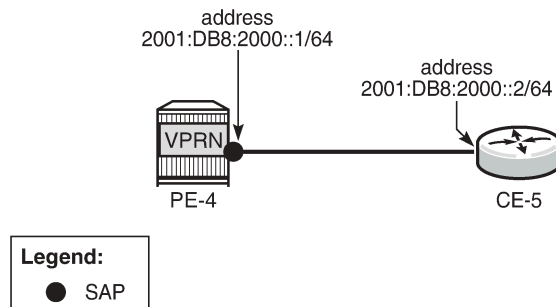
```
# on PE-3
configure
service
  sdp 321 mpls create
  far-end 192.0.2.2
  ldp
  no shutdown
  exit
  vprn 1 name "VPRN 1" customer 1 create
  router-id 192.0.2.31
  autonomous-system 64502
  interface "loopback" create
  address 192.0.2.31/32
  loopback
  exit
  interface "int-PE-3-PE-2" create
  description "Spoke SDP"
  ipv6
  address 2001:db8:1000::2/64
  exit
  spoke-sdp 321:1 create
  no shutdown
  exit
  exit
  bgp-ipvpn
  mpls
  auto-bind-tunnel
  resolution-filter
  ldp
  exit
  resolution filter
  exit
  route-distinguisher 64502:1
  vrf-target target:64502:1
  no shutdown
  exit
  exit
  bgp
  router-id 192.0.2.31
  group "Spoke-CE-1-PE-3"
  family ipv6
  peer-as 64501
  local-address 2001:db8:1000::2
  neighbor 2001:db8:1000::1
  as-override
  type external
  export "PE-3-BGP-CE-1"
  exit
  exit
  no shutdown
  exit
  no shutdown
  exit
  exit
```

In the preceding configuration example, PE-3 has been configured with an IPv6 spoke SDP (spoke termination) with interface int-PE-3-PE-2. The VPRN configuration has also been set up for MP-eBGP peering to CE-1 through the IPv6 spoke interface. The MP-eBGP peering receives and advertises IPv6

prefixes from and to CE-1. The included route policy configuration shows how IPv6 routes are advertised to CE-1 from PE-3 (policy-statement PE-3-BGP-CE-1).

The configuration on PE-4 is similar, but with a SAP interface to CE-5 instead of a spoke-SDP.

Figure 88: PE-4 VPRN with SAP to CE-5



25459

The IPv6 configuration options for the SAP interface (int-PE-4-CE-5) are similar to those in the preceding example for the spoke SDP on PE-3. The PE-4 BGP export policy (PE-4-BGP-CE-5) is also similar to the example for PE-3 in advertising the learned IPv6 route to CE-5.

```
# on PE-4:
configure
  router Base
    policy-options
      begin
      prefix-list "PE-4-CE-5"
        prefix 2001:db8:3000::/64 exact
      exit
      policy-statement "PE-4-BGP-CE-5"
        entry 10
          from
            prefix-list "PE-4-CE-5"
          exit
          action accept
            origin igp
          exit
        exit
        default-action drop
      exit
    exit
  commit
exit
```

```
configure
  service
    vprn 1 name "VPRN 1" customer 1 create
      router-id 192.0.2.41
      autonomous-system 64502
      interface "loopback" create
        address 192.0.2.41/32
        loopback
      exit
      interface "int-PE-4-CE-5" create
        ipv6
          address 2001:db8:2000::1/64
        exit
```

```

        sap 1/1/1 create
        exit
    exit
    bgp-ipvpn
    mpls
        auto-bind-tunnel
        resolution-filter
        ldp
        exit
        resolution filter
    exit
    route-distinguisher 64502:1
    vrf-target target:64502:1
    no shutdown
    exit
exit
exit
bgp
    router-id 192.0.2.41
    group "CE-5-PE-4"
        family ipv6
        peer-as 64501
        local-address 2001:db8:2000::1
        neighbor 2001:db8:2000::2
            as-override
            type external
            export "PE-4-BGP-CE-5"
        exit
    exit
    no shutdown
exit
no shutdown
exit

```

In this setup, the configuration on CE-1 is as follows.

```

# on CE-1:
configure
    router Base
        static-route 2001:db8:3000::/64
            black-hole
            no shutdown
        exit
    exit
    policy-options
        begin
        prefix-list "CE-1-192.0.2.1"
            prefix 2001:db8:3000::/64 exact
        exit
        policy-statement "CE-1-sys-to-eBGP"
            entry 10
                from
                    prefix-list "CE-1-192.0.2.1"
                exit
                action accept
                origin igp
            exit
            default-action drop
        exit
    exit
    commit
exit
bgp

```

```

router-id 192.0.2.1
group "eBGP_to_64502"
  description "eBGP_to_PE-3_AS64502"
  family ipv6
  type external
  peer-as 64502
  local-address 2001:db8:1000::1
  neighbor 2001:db8:1000::2
    export "CE-1-sys-to-eBGP"
  exit
exit
no shutdown
exit
exit
exit

```

```

# on CE-1:
configure
  service
    ies 1 name "IES 1" customer 1 create
    interface "int-CE-1-PE-2" create
      description "SAP_toward_VPRN_Service"
      ipv6
        address 2001:db8:1000::1/64
      exit
      sap 1/1/1 create
    exit
  exit
no shutdown
exit

```

The configuration on CE-5 is similar.

The following command on PE-2 shows that the Epipe VLL is established with the SAP facing CE-1 and spoke SDP facing VPRN 1 on PE-3.

```

*A:PE-2# show service id 1 base
=====
Service Basic Information
=====
Service Id       : 1                Vpn Id          : 0
Service Type    : Epipe
MACSec enabled  : no
Name            : Epipe 1
Description     : (Not Specified)
Customer Id     : 1                Creation Origin  : manual
Last Status Change: 01/20/2022 13:57:46
Last Mgmt Change  : 01/20/2022 13:57:32
Test Service    : No
Admin State     : Up                Oper State      : Up
MTU             : 9190
Vc Switching    : False
SAP Count       : 1                SDP Bind Count  : 1
Per Svc Hashing : 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
<b>sap:1/1/2</b>	<b>null</b>	<b>9212</b>	<b>9212</b>	<b>Up</b>	<b>Up</b>
<b>sdp:231:1 S(192.0.2.3)</b>	<b>Spok</b>	<b>0</b>	<b>9190</b>	<b>Up</b>	<b>Up</b>

The same command can be launched on PE-3 to verify that the VPRN service is up and that the spoke SDP is up (admin state up/oper state up).

```
*A:PE-3# show service id 1 base
=====
Service Basic Information
=====
Service Id       : 1                Vpn Id          : 0
Service Type   : VPRN
MACSec enabled  : no
Name            : VPRN 1
Description     : (Not Specified)
Customer Id     : 1                Creation Origin  : manual
Last Status Change: 01/20/2022 13:57:44
Last Mgmt Change : 01/20/2022 13:57:44
Admin State   : Up              Oper State    : Up

Router Oper State : Up
Route Dist.      : 64502:1         VPRN Type       : regular
Oper Route Dist  : 64502:1
Oper RD Type     : configured
AS Number       : 64502           Router Id        : 192.0.2.31
ECMP             : Enabled         ECMP Max Routes  : 1
Max IPv4 Routes  : No Limit
Local Rt Domain-Id: None         D-Path Lng Ignore : Disabled

Auto Bind Tunnel
Allow Flex-Alg-Fb : Disabled
Resolution        : filter
Filter Protocol   : ldp
Weighted ECMP    : Disabled       ECMP Max Routes  : 1
Strict Tnl Tag    : Disabled

Max IPv6 Routes  : No Limit
Ignore NH Metric : Disabled
Hash Label       : Disabled
Entropy Label    : Disabled
Vrf Target       : target:64502:1
Vrf Import       : None
Vrf Export       : None
MVPN Vrf Target  : None
MVPN Vrf Import  : None
MVPN Vrf Export  : None
Car. Sup C-VPN   : Disabled
Label mode       : vrf
BGP VPN Backup   : Disabled
BGP Export Inactv : Disabled
LOG all events   : Disabled

SAP Count        : 0                SDP Bind Count   : 1
VSD Domain       : <none>

-----
Service Access & Destination Points
-----
Identifier       Type           AdmMTU  OprMTU  Adm  Opr
```

```
-----
sdp:321:1 S(192.0.2.2)                TLDP                0          9190        Up        Up
=====
```

The following command shows that the IPv6 interface is established and its IPv6 address is preferred (2001:db8:1000::2/64). The IPv6 link local address (fe80::17:ffff:fe00:0/64) has been dynamically assigned and is in the preferred state.

```
*A:PE-3# show service id 1 interface
=====
Interface Table
=====
Interface-Name      Adm      Opr(v4/v6)  Type      Port/SapId
IP-Address          PfxState
-----
loopback            Up       Up/Down     VPRN      loopback
192.0.2.31/32      n/a
int-PE-3-PE-2     Up     Down/Up   VPRN     spoke-321:1
2001:db8:1000::2/64      PREFERRED
fe80::17:ffff:fe00:0/64 PREFERRED
-----
Interfaces : 2
=====
```

With the following command, an extensive list of parameters is displayed, including IPv6-related fields that can be checked if configured: DHCP6-relay, DHCP6-server, and so on. It is possible to use filters to reduce the output.

```
*A:PE-3# show service id 1 all
```

After verification of the services (Epipe, VPRN), the MP-eBGP peering connectivity (through IPv6 interfaces) on the VPRN between PE-3 and CE-1 can be verified as follows:

```
*A:PE-3# show router 1 bgp neighbor
=====
BGP Neighbor
=====
-----
Peer          : 2001:db8:1000::1
Description   : (Not Specified)
Group         : Spoke-CE-1-PE-3
-----
Peer AS       : 64501           Peer Port       : 179
Peer Address  : 2001:db8:1000::1
Local AS      : 64502           Local Port      : 49566
Local Address : 2001:db8:1000::2
Peer Type     : External       Dynamic Peer    : No
State       : Established      Last State     : Active
Last Event    : recvOpen
Last Error    : Unrecognized Error
Local Family : IPv6
Remote Family : IPv6
---snip---

Local Capability : RtRefresh MPBGP 4byte ASN
Remote Capability : RtRefresh MPBGP 4byte ASN
Local AddPath Capabi*: Disabled
Remote AddPath Capab*: Send - None
                    : Receive - None
```

```

Import Policy      : None Specified - Default Accept
Export Policy     : PE-3-BGP-CE-1
                  : Default Accept
---snip---

-----
Ingress prefix counters per family.
IPv4 received     : 0
IPv4 active       : 0
IPv4 suppressed   : 0
IPv4 rejected     : 0
VPN-IPv4 received : 0
VPN-IPv4 active   : 0
VPN-IPv4 suppressed : 0
VPN-IPv4 rejected : 0
IPv6 received     : 1
IPv6 active       : 1
IPv6 suppressed   : 0
IPv6 rejected     : 0
VPN-IPv6 received : 0
VPN-IPv6 active   : 0
VPN-IPv6 suppressed : 0
VPN-IPv6 rejected : 0
---snip---

```

Not only is the MP-eBGP session on the VPRN established, but the MP-BGP capabilities are also supported (locally and remotely). PE-3 and its BGP peer CE-1 have advertised and received an IPv6 prefix.

The same command can be launched on PE-4. The status of the VPRN service on PE-4 and its interface to CE-5 can be verified as follows:

```

*A:PE-4# show service id 1 base

=====
Service Basic Information
=====
Service Id       : 1                Vpn Id          : 0
Service Type   : VPRN
MACSec enabled   : no
Name             : VPRN 1
Description      : (Not Specified)
Customer Id      : 1                Creation Origin  : manual
Last Status Change: 01/20/2022 13:57:56
Last Mgmt Change  : 01/20/2022 13:57:56
Admin State   : Up                Oper State     : Up

Router Oper State : Up
Route Dist.       : 64502:1         VPRN Type       : regular
Oper Route Dist   : 64502:1
Oper RD Type      : configured
AS Number        : 64502            Router Id        : 192.0.2.41
ECMP              : Enabled         ECMP Max Routes  : 1
Max IPv4 Routes   : No Limit
Local Rt Domain-Id: None           D-Path Lng Ignore : Disabled

Auto Bind Tunnel
Allow Flex-Alg-Fb : Disabled
Resolution        : filter
Filter Protocol   : ldp
Weighted ECMP     : Disabled         ECMP Max Routes  : 1
Strict Tnl Tag    : Disabled

Max IPv6 Routes   : No Limit
Ignore NH Metric  : Disabled
Hash Label        : Disabled
Entropy Label     : Disabled
Vrf Target        : target:64502:1
Vrf Import        : None
Vrf Export        : None
MVPN Vrf Target   : None

```

```
MVPN Vrf Import   : None
MVPN Vrf Export   : None
Car. Sup C-VPN    : Disabled
Label mode       : vrf
BGP VPN Backup    : Disabled
BGP Export Inactv : Disabled
LOG all events    : Disabled

SAP Count        : 1
VSD Domain       : <none>

SDP Bind Count   : 0
```

-----  
Service Access & Destination Points  
-----

Identifier	Type	AdmMTU	OprMTU	Adm	Opr
<b>sap:1/1/1</b>	<b>null</b>	<b>9212</b>	<b>9212</b>	<b>Up</b>	<b>Up</b>

=====

The VPRN service is up and the SAP is up.

The following command shows that the IPv6 interface is established and its IPv6 address is in the preferred state.

```
*A:PE-4# show service id 1 interface

=====
Interface Table
=====
Interface-Name      Adm    Opr(v4/v6)  Type    Port/SapId
IP-Address          PfxState
-----
loopback            Up     Up/Down     VPRN    loopback
 192.0.2.41/32      n/a
int-PE-4-CE-5     Up    Down/Up    VPRN    1/1/1
  2001:db8:2000::1/64      PREFERRED
  fe80::1b:ffff:fe00:0/64  PREFERRED
-----
Interfaces : 2
=====
```

MP-iBGP, providing 6VPE is configured and built between PE-3 and PE-4 across the MPLS network. IPv6 prefixes are received on PE-3 from CE-1 (2001:db8:3000::/64) and on PE-4 from CE-5 (2001:db8:4000::/64) across the MPLS network using MP-iBGP (6VPE).

CE-1 advertises IPv6 prefix 2001:db8:3000::/64 and CE-5 advertises IPv6 prefix 2001:db8:4000::/64.

The following command on PE-3 shows whether VPN-IPv6 routes were received from and advertised to its iBGP peer PE-4:

```
*A:PE-3# show router bgp summary

=====
BGP Router ID:192.0.2.3      AS:64502      Local AS:64502
=====
BGP Admin State      : Up          BGP Oper State      : Up
Total Peer Groups    : 1            Total Peers         : 1
Total VPN Peer Groups : 1            Total VPN Peers     : 1
Current Internal Groups : 1          Max Internal Groups : 1
Total BGP Paths      : 23           Total Path Memory   : 8168
---snip---

Total VPN-IPv4 Rem. Rts : 0          Total VPN-IPv4 Rem. Act. Rts: 0
```

```

Total VPN-IPv6 Rem. Rts : 2      Total VPN-IPv6 Rem. Act. Rts: 2
Total VPN-IPv4 Bkup Rts : 0      Total VPN-IPv6 Bkup Rts    : 0
Total VPN Local Rts    : 5      Total VPN Supp. Rts       : 0
Total VPN Hist. Rts   : 0      Total VPN Decay Rts      : 0
---snip---

```

```

=====
BGP Summary
=====
Legend : D - Dynamic Neighbor
=====
Neighbor
Description
          AS PktRcvd InQ Up/Down  State|Rcv/Act/Sent (Addr Family)
          PktSent OutQ
-----
192.0.2.4
PE-4
          64502      38      0 00h16m24s 2/2/2 (VpnIPv6)
          40      0
-----

```

PE-3 has received and learned a valid and best IPv6 route for prefix 2001:db8:3000::/64 with a BGP next hop of 2001:db8:1000:: (CE-1), as follows:

```

*A:PE-3# show router 1 bgp routes ipv6
=====
BGP Router ID:192.0.2.31      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
Origin codes : i - IGP, e - EGP, ? - incomplete
=====
BGP IPv6 Routes
=====
Flag Network                LocalPref MED
      Nexthop (Router)      Path-Id   IGP Cost
      As-Path                Label
-----
u*>i 2001:db8:3000::/64        None     None
      2001:db8:1000::1      None     0
      64501                  -
-----
Routes : 1
=====

```

The following output shows the 2001:db8:4000::/64 prefix as BGP IPv6 route advertised by PE-3 to eBGP peer CE-1.

```

*A:PE-3# show router 1 bgp neighbor 2001:db8:1000::1 advertised-routes ipv6
=====
BGP Router ID:192.0.2.31      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
Origin codes : i - IGP, e - EGP, ? - incomplete
=====
BGP IPv6 Routes

```

```

=====
Flag Network                               LocalPref MED
  Nexthop (Router)                         Path-Id   IGP Cost
  As-Path                                     Label
-----
i   2001:db8:4000::/64                       n/a      None
    2001:db8:1000::2                          None     n/a
    64502 64502                                -        -
-----
Routes : 1
=====

```

The IPv6 route 2001:db8:4000::/64 originates from CE-5 and was advertised from CE-5 to its eBGP peer PE-4, then from PE-4 as VPN-IPv6 route to its iBGP peer PE-3 with next-hop PE-4, as follows:

```

*A:PE-3# show router bgp routes vpn-ipv6
=====
BGP Router ID:192.0.2.3      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP VPN-IPv6 Routes
=====
Flag Network                               LocalPref MED
  Nexthop (Router)                         Path-Id   IGP Cost
  As-Path                                     Label
-----
u*>i 64502:1:2001:db8:2000::/64              100      None
     ::ffff:192.0.2.4                          None     10
     No As-Path                                524284
u*>i 64502:1:2001:db8:4000::/64             100      None
     ::ffff:192.0.2.4                          None     10
     64501                                       524284
-----
Routes : 2
=====

```

PE-3 is advertising the VPN-IPv6 route of 2001:db8:3000::/64 to its MP-iBGP peer PE-4, as follows. The IPv6 prefix 2001:db8:3000::/64 was learned from CE-1 in an MP-eBGP session:

```

*A:PE-3# show router bgp neighbor 192.0.2.4 advertised-routes vpn-ipv6
=====
BGP Router ID:192.0.2.3      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP VPN-IPv6 Routes
=====
Flag Network                               LocalPref MED
  Nexthop (Router)                         Path-Id   IGP Cost
  As-Path                                     Label
-----
i   64502:1:2001:db8:1000::/64              100      None
     ::ffff:192.0.2.3                          None     n/a
-----

```

```

i      No As-Path
      64502:1:2001:db8:3000::/64
      ::ffff:192.0.2.3
      64501
      100
      None
      524284
      None
      n/a
      524284
-----
Routes : 2
=====

```

The list of VPN-IPv6 routes on PE-4 includes the VPN-IPv6 route that was learned from PE-3: 2001:db8:3000::/64, as follows:

```

*A:PE-4# show router bgp routes vpn-ipv6
=====
BGP Router ID:192.0.2.4      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP VPN-IPv6 Routes
=====
Flag  Network                               LocalPref  MED
      Nexthop (Router)                     Path-Id    IGP Cost
      As-Path                               Label
-----
u*>i  64502:1:2001:db8:1000::/64
      ::ffff:192.0.2.3
      No As-Path
      100
      None
      10
      524284
u*>i  64502:1:2001:db8:3000::/64
      ::ffff:192.0.2.3
      64501
      100
      None
      10
      524284
-----
Routes : 2
=====

```

The following output shows the advertised VPN-IPv6 route of 2001:db8:4000::/64 from PE-4 to PE-3.

```

*A:PE-4# show router bgp neighbor 192.0.2.3 advertised-routes vpn-ipv6
=====
BGP Router ID:192.0.2.4      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP VPN-IPv6 Routes
=====
Flag  Network                               LocalPref  MED
      Nexthop (Router)                     Path-Id    IGP Cost
      As-Path                               Label
-----
i      64502:1:2001:db8:2000::/64
      ::ffff:192.0.2.4
      No As-Path
      100
      None
      n/a
      524284
i      64502:1:2001:db8:4000::/64
      ::ffff:192.0.2.4
      64501
      100
      None
      n/a
      524284
-----

```

```
Routes : 2
```

The following output from PE-4 shows the IPv6 prefix 2001:db8:4000::/64 learned from CE-5.

```
*A:PE-4# show router 1 bgp routes ipv6
=====
BGP Router ID:192.0.2.41      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP IPv6 Routes
=====
Flag Network                               LocalPref MED
  Nexthop (Router)                         Path-Id   IGP Cost
  As-Path                                   Label
-----
u*>i 2001:db8:4000::/64                      None     None
      2001:db8:2000::2                       None     0
      64501                                    -
-----
Routes : 1
=====
```

The following command on PE-4 confirms that IPv6 prefix 2001:db8:3000::/64 is advertised to CE-5.

```
*A:PE-4# show router 1 bgp neighbor 2001:db8:2000::2 advertised-routes ipv6
=====
BGP Router ID:192.0.2.41      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP IPv6 Routes
=====
Flag Network                               LocalPref MED
  Nexthop (Router)                         Path-Id   IGP Cost
  As-Path                                   Label
-----
i    2001:db8:3000::/64                      n/a      None
      2001:db8:2000::1                       None     n/a
      64502 64502                             -
-----
Routes : 1
=====
```

The final verification of CE-1 and CE-5 shows that IPv6 routes for AS 64501 have been received and are valid across the VPRN service, as follows:

```
*A:CE-1# show router bgp routes ipv6
=====
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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP IPv6 Routes
=====
Flag Network                               LocalPref MED
  Nextthop (Router)                       Path-Id   IGP Cost
  As-Path                                  Label
-----
u*>i 2001:db8:4000::/64                       None     None
      2001:db8:1000::2                       None     0
      64502 64502                             -
-----
Routes : 1
=====

```

```

*A:CE-5# show router bgp routes ipv6
=====
BGP Router ID:192.0.2.5      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
Origin codes  : i - IGP, e - EGP, ? - incomplete
=====
BGP IPv6 Routes
=====
Flag Network                               LocalPref MED
  Nextthop (Router)                       Path-Id   IGP Cost
  As-Path                                  Label
-----
u*>i 2001:db8:3000::/64                       None     None
      2001:db8:2000::1                       None     0
      64502 64502                             -
-----
Routes : 1
=====

```

## Conclusion

Spoke termination for IPv6-6VPE extends the use of spoke terminated interfaces from an Epipe VLL into a VPRN service using IPv6 interfaces on the access. Supporting the requirement of IPv6 interfaces, routing of IPv6 prefixes and the use of 6VPE for IPv6 tunneling over an IPv4 network allows SR OS to provide capabilities supporting the growth of IPv6 architectures.

# Traffic Leaking from VPRN to GRT

This chapter provides information about Traffic Leaking from VPRN to GRT.

Topics in this chapter include:

- [Applicability](#)
- [Overview](#)
- [Configuration](#)
- [Conclusion](#)

## Applicability

The information and configuration in this chapter were originally based on SR OS Release 14.0 R4. The CLI in the current edition corresponds to SR OS Release 22.2.R2.

## Overview

RFC 4364, *BGP/MPLS IP Virtual Private Networks (VPNs)*, describes a method of distributing routing information using BGP and MPLS forwarding data to provide a Layer 3 VPN service to end users. Each Virtual Private Routed Network (VPRN) consists of a set of customer sites connected to one or more PE routers. Each associated PE router maintains a separate IP forwarding table for each VPRN. Additionally, the PE routers exchange the routing information configured or learned from all customer sites via Multi-Protocol Border Gateway Protocol (MP-BGP) peering. Each route exchanged via the MP-BGP protocol includes a route distinguisher (RD), which identifies the VPRN association and resolves any IP address overlap.

It has always been possible to exchange traffic from one VPRN to another, using scenarios such as "extranet", "hub and spoke" and so on, using the vrf-import and vrf-export policies for BGP VPN-IPv4 route distribution.

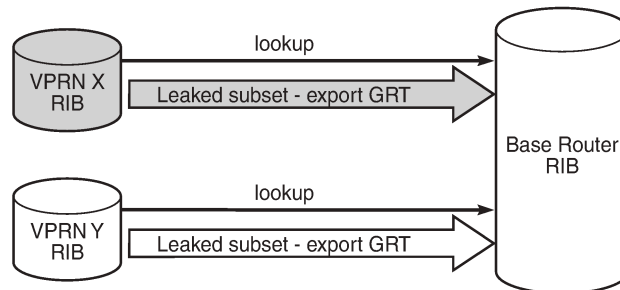
Traffic leaking to the Global Route Table (GRT) allows service providers to offer VPRN and Internet services over a single virtual routing and forwarding VRF interface. Packets entering a VRF interface can have route processing results derived from the VRF or the GRT. The leaking and preferred lookup settings are configured on a per-VPRN basis.

To allow data flowing from a VPRN to the base router, routing information from the base router must be made available for lookup by the VPRN. The GRT lookup can be general (for example, any lookup miss in the Virtual Routing and Forwarding (VRF) table can be resolved in the GRT), or specific (for example, specific routes should only be searched for in the GRT and ignored by the VPRN).

To enable the GRT lookup from the VPRN, the **enable-grt** command is used. This only provides part of the solution, because packets can now be forwarded from the VPRN to the GRT, but not in the opposite direction. The GRT needs to learn specific destination prefixes from the VPRN and this is achieved by route leaking from the VPRN to the GRT, using policies (**export-grt** command). The maximum number of routes leaked from a VPRN to the GRT is five by default, but this maximum can be modified or

even removed. Prefixes should be globally unique within the service provider network and if these are propagated outside the provider's network, they must be from the public IP space and globally unique.

Figure 89: VPRN to GRT leak process



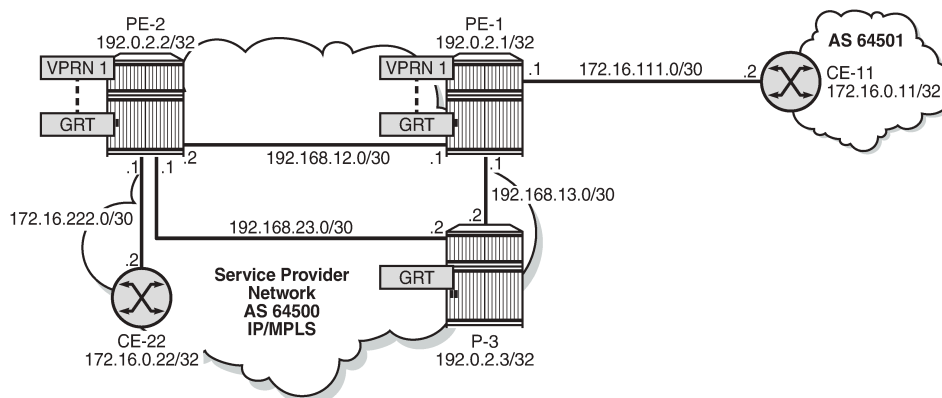
25998

The method described in this chapter allows the network administrator to leak specific or all routes that are inside a VPRN to the GRT. Route leaking from VPRN to GRT is protocol-independent and can be applied for BGP, OSPF(v3), IS-IS, static, local routes, and so on. For BGP routes, there is an improved route leaking mechanism that allows leaking routes preserving all BGP attributes; see chapter *BGP Route Leaking*.

## Configuration

Figure 90: Example topology with IPv4 addresses shows the example topology used in this chapter, including the IPv4 addresses. The interfaces also have IPv6 addresses, which will be shown in Figure 92: Example topology with IPv6 addresses.

Figure 90: Example topology with IPv4 addresses



25999

## Initial configuration

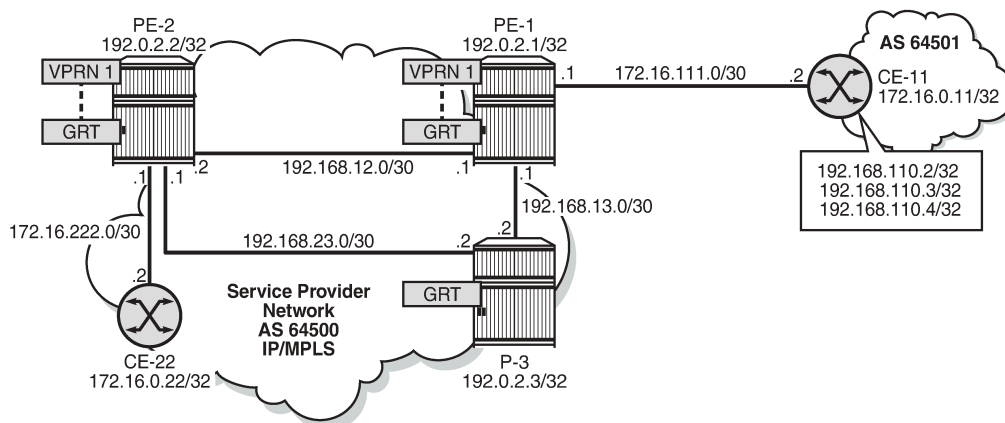
The nodes in the example topology have the following initial configuration:

- Cards, MDAs, ports
- Router interfaces
- IGP (IS-IS or OSPF) between the PEs
- LDP between the PEs
- VPRN "VPRN 1" on PE-1
- BGP (IBGP between the PEs; EBGP between PE-1 and CE-11)
  - On PE-1, BGP is configured in the base router and in VPRN 1.
- Loopback addresses on CE-11, such as 192.168.110.2/32.
- Export policies on CE-11 to export routes from direct with certain prefixes.

### Protocol-independent IPv4 route leaking from VPRN to GRT

Figure 91: IPv4 VPRN to GRT route leaking for IS-IS shows the topology with the IP addresses for this example. Route leaking from VPRN to GRT is protocol independent and in this example, VPRN "VPRN 1" on PE-1 will leak local routes, static routes, and imported BGP routes to the GRT. IS-IS or OSPF routes can also be leaked, but that is not shown here.

Figure 91: IPv4 VPRN to GRT route leaking for IS-IS



26000

GRT-leak is by default disabled. The routing table for VPRN 1 on PE-1 contains local routes, static routes, and BGP routes that are learned from CE-11, as follows:

```
*A: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  00h01m14s    0
system                                           0
172.16.111.0/30                                   Local  Local  00h01m14s    0
int-PE-1-CE-11                                  0
```

```

192.168.110.2/32          Remote BGP      00h00m12s 170
    172.16.111.2          0
192.168.110.3/32          Remote BGP      00h00m12s 170
    172.16.111.2          0
192.168.110.4/32          Remote BGP      00h00m12s 170
    172.16.111.2          0
192.168.120.0/24         Remote Static   00h01m14s 5
    172.16.111.2          1
-----
No. of Routes: 6
Flags: n = Number of times nexthop is repeated
      B = BGP backup route available
      L = LFA nexthop available
      S = Sticky ECMP requested
=====

```

By default, the GRT is not learning the VPRN routes, as follows:

```

*A:PE-1# show router route-table

=====
Route Table (Router: Base)
=====
Dest Prefix[Flags]          Type  Proto  Age      Pref
  Next Hop[Interface Name]  Metric
-----
192.0.2.1/32                Local  Local  00h01m14s 0
    system
192.0.2.2/32                Remote  ISIS   00h01m03s 15
    192.168.12.2          10
192.0.2.3/32                Remote  ISIS   00h00m51s 15
    192.168.13.2          10
192.168.12.0/30             Local  Local  00h01m14s 0
    int-PE-1-PE-2        0
192.168.13.0/30             Local  Local  00h01m14s 0
    int-PE-1-P-3         0
192.168.23.0/30             Remote  ISIS   00h01m03s 15
    192.168.12.2          20
-----
No. of Routes: 6
Flags: n = Number of times nexthop is repeated
      B = BGP backup route available
      L = LFA nexthop available
      S = Sticky ECMP requested
=====

```

To enable VPRN to GRT leaking, the following route policy is configured on PE-1 and applied in VPRN 1:

```

# on PE-1:
configure
  router Base
    policy-options
      begin
        policy-statement "LeakVPRNtoGRT_pref8"
          entry 10
            action accept
            preference 8
          exit
        exit
      exit
    exit
  commit
exit
exit

```

```

service
  vprn "VPRN 1"
  grt-lookup
  enable-grt
  exit
  export-grt "LeakVPRNtoGRT_pref8"
  exit
exit

```

This policy allows leaking all routes from a VPRN to the base router, without any match criteria. However, when routes are leaked from VPRNs to the GRT, they need to be unique and only routes that need to be known in the GRT should be leaked. By default, the preference for a leaked route is 180. The preference can be manually configured to a lower value, such as 8, to avoid network inconsistencies between the IGP and the RT on the router where the routes are leaked.

When **enable-grt** is configured, any lookup miss in the VRF table will be resolved in the GRT, if available. This only works from VPRN to GRT and does not require route leaking. However, the base router needs to be able to route packets back to the VPRN and it cannot perform a lookup in the routing table of the VPRN. Therefore, route leaking from VPRN to GRT is required, and **export-grt** is configured. Prefixes in the VPRN must be leaked to the GRT through a policy. Prefixes leaked from any VPRN should never conflict with prefixes leaked from any other VPRN or existing prefixes in the GRT.

This configuration is protocol-independent. Route leaking from VPRN to GRT is applicable for all kinds of learned routes, such as static routes, local routes, IS-IS, OSPF, BGP, and so on.

After routes are leaked from the VPRN to the GRT, the routing table of the base router includes the leaked routes, with protocol "VPN Leak". For PE-1, the routing table contains the following routes:

```

*A:PE-1# show router route-table
=====
Route Table (Router: Base)
=====
Dest Prefix[Flags]                Type  Proto  Age           Pref
  Next Hop[Interface Name]                Metric
-----
172.16.1.1/32                      Remote VPN Leak 00h00m20s  8
  system                               0
172.16.111.0/30                     Remote VPN Leak 00h00m20s  8
  int-PE-1-CE-11                       0
192.0.2.1/32                        Local  Local  00h03m06s    0
  system                               0
192.0.2.2/32                        Remote  ISIS  00h02m49s   15
  192.168.12.2                          10
192.0.2.3/32                        Remote  ISIS  00h02m41s   15
  192.168.13.2                          10
192.168.12.0/30                     Local  Local  00h03m06s    0
  int-PE-1-PE-2                         0
192.168.13.0/30                     Local  Local  00h03m06s    0
  int-PE-1-P-3                          0
192.168.23.0/30                     Remote  ISIS  00h02m49s   15
  192.168.12.2                          20
192.168.110.3/32                    Remote VPN Leak 00h00m20s  8
  172.16.111.2                          0
192.168.110.4/32                    Remote VPN Leak 00h00m20s  8
  172.16.111.2                          0
192.168.120.0/24                    Remote VPN Leak 00h00m20s  8
  172.16.111.2                          0
-----
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
```

Regardless the preference of the original routes in VPRN 1, all the leaked routes in the GRT have preference 8, as configured. By default, a maximum of five routes are leaked. This export limit can be overruled, as follows:

```
# on PE-1:
configure
  service
    vprn "VPRN 1"
      grt-lookup
      export-limit 10
```

The following command shows only the routes leaked from any VPRN to GRT on PE-1:

```
*A:PE-1# show router route-table protocol vpn-leak all
```

```
=====
Route Table (Router: Base)
=====
```

Dest Prefix[Flags] Next Hop[Interface Name]	Type	Proto Active	Age Metric	Pref
172.16.1.1/32 system	Remote	VPN Leak Y	00h00m05s 0	8
172.16.111.0/30 int-PE-1-CE-11	Remote	VPN Leak Y	00h00m05s 0	8
192.168.110.2/32 172.16.111.2	Remote	VPN Leak Y	00h00m05s 0	8
192.168.110.3/32 172.16.111.2	Remote	VPN Leak Y	00h00m05s 0	8
192.168.110.4/32 172.16.111.2	Remote	VPN Leak Y	00h00m05s 0	8
192.168.120.0/24 172.16.111.2	Remote	VPN Leak Y	00h00m05s 0	8

```
-----
No. of Routes: 6
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
       E = Inactive best-external BGP route
=====
```

Different types of routes are leaked to the GRT with protocol type "VPN Leak" and all of them get the same preference, configured or default. The detailed output for any leaked route in the preceding list for PE-1 shows protocol VPN\_LEAK and preference 8, as follows:

```
*A:PE-1# show router route-table protocol vpn-leak 192.168.110.2/32 extensive
```

```
=====
Route Table (Router: Base)
=====
```

Dest Prefix	: 192.168.110.2/32
<b>Protocol</b>	: <b>VPN_LEAK</b>
Age	: 00h00m24s
<b>Preference</b>	: <b>8</b>
Next-Hop	: 172.16.111.2

```

Interface      : int-PE-1-CE-11 (VPRN 1)
QoS           : Priority=n/c, FC=n/c
Source-Class  : 0
Dest-Class    : 0
Metric        : 0
ECMP-Weight   : N/A
-----
No. of Destinations: 1
=====

```

## Export IPv4 VPN-leak routes to routing protocols

Until now, the VPN-leak routes are leaked locally to the GRT, but they are not advertised in IS-IS, OSPF, or BGP. Router P-3 has not learned any of the leaked routes, as follows:

```

*A:P-3# show router route-table

=====
Route Table (Router: Base)
=====
Dest Prefix[Flags]          Type  Proto  Age      Pref
  Next Hop[Interface Name]  Metric
-----
192.0.2.1/32                Remote  ISIS   00h03m44s  15
    192.168.13.1             10
192.0.2.2/32                Remote  ISIS   00h03m44s  15
    192.168.23.1             10
192.0.2.3/32                Local   Local  00h03m51s  0
    system                    0
192.168.12.0/30             Remote  ISIS   00h03m44s  15
    192.168.13.1             20
192.168.13.0/30             Local   Local  00h03m51s  0
    int-P-3-PE-1              0
192.168.23.0/30             Local   Local  00h03m51s  0
    int-P-3-PE-2              0
-----
No. of Routes: 6
Flags: n = Number of times nexthop is repeated
      B = BGP backup route available
      L = LFA nexthop available
      S = Sticky ECMP requested
=====

```

To reduce the number of routes to be exported on PE-1, a match criterion is added for the routes to be leaked, as follows:

```

# on PE-1:
configure
  router Base
    policy-options
      begin
      prefix-list "192.168.110.0"
        prefix 192.168.110.0/24 longer
      exit
      policy-statement "LeakVPRNtoGRT_pref8_110"
        entry 10
          from
            prefix-list "192.168.110.0"
          exit
        action accept

```



```

                preference 8
            exit
        exit
    exit
    commit
exit
service
  vprn "VPRN 1"
    grt-lookup
      enable-grt
    exit
    export-grt "LeakVPRNtoGRT_pref8_110"
  exit

```

VPN-leak routes can be exported to any routing protocol. Prefix lists can be used to filter routes, but that is not configured in this example. The following export policy is configured on PE-1 to export the VPN-leak routes:

```

# on PE-1:
configure
  router Base
    policy-options
      begin
        policy-statement "export-vpn-leak"
          entry 10
            from
              protocol vpn-leak
            exit
            action accept
          exit
        exit
      exit
    exit
  commit

```

The same export policy will be used for export to IS-IS, OSPF, and BGP.

## Export IPv4 VPN-leak routes to IS-IS

The export policy is applied in the IS-IS context on PE-1, as follows:

```

# on PE-1:
configure
  router Base
    isis 0
      export "export-vpn-leak"
    exit

```

The leaked routes are now advertised via IS-IS and appear as IS-IS routes with default preference for IS-IS routes on PE-2 and P-3. The route table on P-3 looks as follows:

```

*A:P-3# show router route-table

=====
Route Table (Router: Base)
=====
Dest Prefix[Flags]                                Type   Proto   Age      Pref
  Next Hop[Interface Name]                        Metric
-----

```

```

192.0.2.1/32          Remote  ISIS    00h04m39s  15
    192.168.13.1      10
192.0.2.2/32          Remote  ISIS    00h04m39s  15
    192.168.23.1      10
192.0.2.3/32          Local   Local   00h04m46s  0
    system            0
192.168.12.0/30       Remote  ISIS    00h04m39s  15
    192.168.13.1      20
192.168.13.0/30       Local   Local   00h04m46s  0
    int-P-3-PE-1      0
192.168.23.0/30       Local   Local   00h04m46s  0
    int-P-3-PE-2      0
192.168.110.2/32      Remote  ISIS    00h00m21s  15
    192.168.13.1      10
192.168.110.3/32      Remote  ISIS    00h00m21s  15
    192.168.13.1      10
192.168.110.4/32      Remote  ISIS    00h00m21s  15
    192.168.13.1      10
-----

```

```

No. of Routes: 9
Flags: n = Number of times nexthop is repeated
      B = BGP backup route available
      L = LFA nexthop available
      S = Sticky ECMP requested
=====

```

The export policy is removed from the IS-IS context on PE-1, as follows:

```

# on PE-1:
configure
router Base
isis 0
no export

```

## Export IPv4 VPN-leak routes to OSPF

When OSPF is used instead of IS-IS, the behavior is similar. The export policy is applied in the OSPF context on PE-1, as follows:

```

# on PE-1:
configure
router Base
ospf 0
export "export-vpn-leak"

```

To export routes into OSPF using a policy, the router must be configured as ASBR, as follows:

```

# on PE-1:
configure
router Base
ospf 0
asbr

```

The routes with protocol VPN-leak on PE-1 are now exported in OSPF to PE-2 and P-3. The default preference for external OSPF routes is 150. On P-3, the routing table contains the following OSPF routes:

```

*A:P-3# show router route-table protocol ospf

```

```

=====
Route Table (Router: Base)
=====
Dest Prefix[Flags]                Type  Proto  Age           Pref
  Next Hop[Interface Name]          Metric
-----
192.0.2.1/32                      Remote OSPF   00h00m44s    10
    192.168.13.1                    10
192.0.2.2/32                      Remote OSPF   00h00m44s    10
    192.168.23.1                    10
192.168.12.0/30                   Remote OSPF   00h00m44s    10
    192.168.13.1                    20
192.168.110.2/32                 Remote OSPF   00h00m14s  150
    192.168.13.1                  1
192.168.110.3/32                 Remote OSPF   00h00m14s  150
    192.168.13.1                  1
192.168.110.4/32                 Remote OSPF   00h00m14s  150
    192.168.13.1                  1
-----
No. of Routes: 6
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
=====

```

The export policy is removed from the OSPF context on PE-1 as follows:

```

# on PE-1:
configure
router Base
  ospf
    no export

```

## Export IPv4 VPN-leak routes to BGP

The export policy is applied in the general **bgp** context of PE-1, as follows:

```

# on PE-1:
configure
router Base
  bgp
    export "export-vpn-leak"

```

The VPN-leak routes from PE-1 will be advertised as BGP routes to BGP neighbors PE-2 and P-3, and the routing tables will contain BGP routes with preference 170. P-3 has the following BGP routes:

```

*A:P-3# show router route-table protocol bgp
=====
Route Table (Router: Base)
=====
Dest Prefix[Flags]                Type  Proto  Age           Pref
  Next Hop[Interface Name]          Metric
-----
192.168.110.2/32                 Remote BGP   00h00m16s  170
    192.168.13.1                  10
192.168.110.3/32                 Remote BGP   00h00m16s  170
    192.168.13.1                  10

```

```

192.168.110.4/32          Remote BGP      00h00m16s 170
192.168.13.1            10
-----
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
=====

```



**Note:**

If it is required to preserve the BGP path attributes in the leaking process, you must use the BGP Route Leaking process described in chapter *BGP Route Leaking*. However, with this protocol-independent route leaking mechanism, it is possible to leak non-BGP routes to the GRT that will be advertised as BGP routes.

The export policy is removed from the **bgp** context, as follows:

```

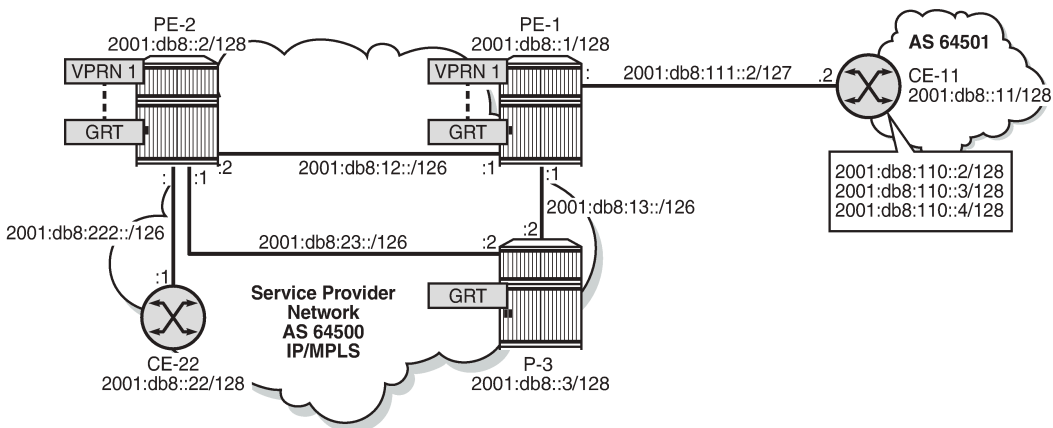
# on PE-1:
configure
router Base
  bgp
    no export

```

### Protocol-independent IPv6 route leaking from VPRN to GRT

Figure 92: Example topology with IPv6 addresses shows the topology and the IP addresses used for IPv6. CE-11 exports routes such as 2001:db8:110::2/128 to VPRN 1 on PE-1. On PE-1, local routes, static routes, and BGP routes will be leaked to the GRT.

Figure 92: Example topology with IPv6 addresses



26001

The IPv6 routing table for VPRN 1 on PE-1 includes local addresses, a static route, and three BGP routes exported by CE-11, as follows:

```

*A:PE-1# show router 1 route-table ipv6
=====

```

```
IPv6 Route Table (Service: 1)
=====
Dest Prefix[Flags]
  Next Hop[Interface Name]
-----
2001:db8::1:1/128
  system
  Local Local 00h08m49s 0
  0
2001:db8:110::2/128
  2001:db8:111::1
  Remote BGP 00h07m49s 170
  0
2001:db8:110::3/128
  2001:db8:111::1
  Remote BGP 00h07m49s 170
  0
2001:db8:110::4/128
  2001:db8:111::1
  Remote BGP 00h07m49s 170
  0
2001:db8:111::/127
  int-PE-1-CE-11
  Local Local 00h08m48s 0
  0
2001:db8:120::/120
  2001:db8:111::1
  Remote Static 00h08m48s 5
  1
-----
No. of Routes: 6
Flags: n = Number of times nexthop is repeated
      B = BGP backup route available
      L = LFA nexthop available
      S = Sticky ECMP requested
=====
```

By default, route leaking is disabled and the IPv6 GRT on PE-1 does not contain any of the IPv6 routes in VPRN 1, as follows:

```
*A:PE-1# show router route-table ipv6
=====
IPv6 Route Table (Router: Base)
=====
Dest Prefix[Flags]
  Next Hop[Interface Name]
-----
2001:db8::1/128
  system
  Local Local 00h08m50s 0
  0
2001:db8::2/128
  fe80::14:1ff:fe01:1-"int-PE-1-PE-2"
  Remote OSPF3 00h08m38s 10
  10
2001:db8::3/128
  fe80::18:1ff:fe01:2-"int-PE-1-P-3"
  Remote OSPF3 00h08m26s 10
  10
2001:db8:12::/126
  int-PE-1-PE-2
  Local Local 00h08m49s 0
  0
2001:db8:13::/126
  int-PE-1-P-3
  Local Local 00h08m49s 0
  0
2001:db8:23::/126
  fe80::18:1ff:fe01:2-"int-PE-1-P-3"
  Remote OSPF3 00h08m22s 10
  20
-----
No. of Routes: 6
Flags: n = Number of times nexthop is repeated
      B = BGP backup route available
      L = LFA nexthop available
      S = Sticky ECMP requested
=====
```

The VPN-leak route policy is the same as for IPv4 routes, and is applied in the **vprn** context in the same way as for IPv4 routes, as follows:

```
# on PE-1:
configure
router Base
```

```

policy-options
  begin
    policy-statement "LeakVPRNtoGRT_pref8"
      entry 10
        action accept
          preference 8
        exit
      exit
    exit
  exit
  commit
exit
service
  vprn "VPRN 1"
    grt-lookup
      enable-grt
    exit
    export-grt "LeakVPRNtoGRT_pref8"
  exit

```

On PE-1, the IPv6 routing table for VPRN 1 contains six routes, but by default, a maximum of five routes are leaked, as follows:

```

*A:PE-1# show router route-table ipv6 protocol vpn-leak all
=====
IPv6 Route Table (Router: Base)
=====
Dest Prefix[Flags]
  Next Hop[Interface Name]
Type      Proto   Age      Pref
          Active  Metric
-----
2001:db8::1:1/128      Remote  VPN Leak 00h00m12s  8
                    system
                    Y
                    0
2001:db8:110::2/128    Remote  VPN Leak 00h00m12s  8
                    2001:db8:111::1
                    Y
                    0
2001:db8:110::4/128    Remote  VPN Leak 00h00m12s  8
                    2001:db8:111::1
                    Y
                    0
2001:db8:111::/127     Remote  VPN Leak 00h00m12s  8
                    int-PE-1-CE-11
                    Y
                    0
2001:db8:120::/120     Remote  VPN Leak 00h00m12s  8
                    2001:db8:111::1
                    Y
                    0
-----
No. of Routes: 5
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
       E = Inactive best-external BGP route
=====

```

The export limit for IPv6 routes is removed, as follows:

```

# on PE-1:
configure
  service
    vprn "VPRN 1"
      grt-lookup
        export-v6-limit 0

```

As a result, there is no limit to the number of leaked IPv6 routes, and all six IPv6 routes are leaked from VPRN 1 to the GRT with the configured preference 8, as follows:

```
*A:PE-1# show router route-table ipv6 protocol vpn-leak all

=====
IPv6 Route Table (Router: Base)
=====
Dest Prefix[Flags]                Type   Proto   Age      Pref
  Next Hop[Interface Name]        Active Metric
-----
2001:db8::1:1/128                 Remote VPN Leak 00h00m05s 8
      system                       Y      0
2001:db8:110::2/128               Remote VPN Leak 00h00m05s 8
      2001:db8:111::1              Y      0
2001:db8:110::3/128               Remote VPN Leak 00h00m05s 8
      2001:db8:111::1              Y      0
2001:db8:110::4/128               Remote VPN Leak 00h00m05s 8
      2001:db8:111::1              Y      0
2001:db8:111::/127                Remote VPN Leak 00h00m05s 8
      int-PE-1-CE-11               Y      0
2001:db8:120::/120                Remote VPN Leak 00h00m05s 8
      2001:db8:111::1              Y      0
-----
No. of Routes: 6
Flags: n = Number of times nexthop is repeated
      B = BGP backup route available
      L = LFA nexthop available
      S = Sticky ECMP requested
      E = Inactive best-external BGP route
=====
```

The details for any of the routes shows that the protocol is VPN-leak and the preference is 8, as follows:

```
*A:PE-1# show router route-table protocol vpn-leak 2001:db8:110::2/128 extensive

=====
Route Table (Router: Base)
=====
Dest Prefix      : 2001:db8:110::2/128
Protocol       : VPN_LEAK
Age              : 00h00m26s
Preference    : 8
Next-Hop        : 2001:db8:111::1
  Interface     : int-PE-1-CE-11 (VPRN 1)
  QoS           : Priority=n/c, FC=n/c
  Source-Class  : 0
  Dest-Class    : 0
  Metric        : 0
  ECMP-Weight   : N/A
-----
No. of Destinations: 1
=====
```

## Export IPv6 VPN-leak routes to routing protocols

Until now, the IPv6 VPN-leak routes are leaked locally to the GRT, but they are not advertised in IS-IS, OSPFv3, or BGP. Router P-3 has not learned any of the leaked IPv6 routes, as follows:

```
*A:P-3# show router route-table ipv6

=====
IPv6 Route Table (Router: Base)
=====
Dest Prefix[Flags]                               Type  Proto  Age           Pref
  Next Hop[Interface Name]                       Metric
-----
2001:db8::1/128                                  Remote OSPF3  00h10m51s  10
      fe80::10:1ff:fe01:1-"int-P-3-PE-1"         10
2001:db8::2/128                                  Remote OSPF3  00h10m46s  10
      fe80::14:1ff:fe01:2-"int-P-3-PE-2"         10
2001:db8::3/128                                  Local  Local  00h10m52s  0
      system                                       0
2001:db8:12::/126                                Remote OSPF3  00h10m51s  10
      fe80::10:1ff:fe01:1-"int-P-3-PE-1"         20
2001:db8:13::/126                                Local  Local  00h10m51s  0
      int-P-3-PE-1                               0
2001:db8:23::/126                                Local  Local  00h10m51s  0
      int-P-3-PE-2                               0
-----
No. of Routes: 6
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
=====
```

To reduce the number of VPN-leak routes, a match criterion is added to the route policy on PE-1, as follows:

```
# on PE-1:
configure
  router Base
    policy-options
      begin
      prefix-list "2001:db8:110::"
        prefix 2001:db8:110::/125 longer
      exit
      policy-statement "LeakVPRNtoGRT_pref8_110"
        entry 20
          from
            prefix-list "2001:db8:110::"
          exit
          action accept
            preference 8
          exit
        exit
      exit
    commit
  exit
exit
service
  vprn "VPRN 1"
    grt-lookup
      enable-grt
```



```

exit
export-grt "LeakVPRNtoGRT_pref8_110"
exit
    
```

The following IPv6 routes are leaked from VPRN 1 to GRT on PE-1:

```

*A:PE-1# show router route-table ipv6 protocol vpn-leak

=====
IPv6 Route Table (Router: Base)
=====
Dest Prefix[Flags]
Next Hop[Interface Name]
Type Proto Age Pref
Metric
-----
2001:db8:110::2/128 Remote VPN Leak 00h00m21s 8
2001:db8:111::1 0
2001:db8:110::3/128 Remote VPN Leak 00h00m21s 8
2001:db8:111::1 0
2001:db8:110::4/128 Remote VPN Leak 00h00m21s 8
2001:db8:111::1 0
-----
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
=====
    
```

IPv6 VPN-leak routes can be exported to routing protocols IS-IS, OSPFv3, and BGP.

The export policy on PE-1 is the same as in all the preceding examples for IPv4, as follows:

```

# on PE-1:
configure
router Base
  policy-options
  begin
  policy-statement "export-vpn-leak"
  entry 10
  from
  protocol vpn-leak
  exit
  action accept
  exit
  exit
exit
commit
    
```

## Export IPv6 VPN-leak routes to IS-IS

The export policy for IPv6 routes of protocol VPN-leak is applied for IS-IS, as follows:

```

# on PE-1:
configure
router Base
  isis 0
  export "export-vpn-leak"
    
```

The three IPv6 VPN-leak routes from PE-1 are now advertised by IS-IS to PE-2 and P-3. The routing table on P-3 contains the following IPv6 IS-IS routes:

```
*A:P-3# show router route-table ipv6 protocol isis

=====
IPv6 Route Table (Router: Base)
=====
Dest Prefix[Flags]                Type  Proto  Age      Pref
Next Hop[Interface Name]         Metric
-----
2001:db8::1/128                   Remote ISIS  00h11m54s 15
                                fe80::10:1ff:fe01:1-"int-P-3-PE-1"
                                10
2001:db8::2/128                   Remote ISIS  00h11m54s 15
                                fe80::14:1ff:fe01:2-"int-P-3-PE-2"
                                10
2001:db8:12::/126                 Remote ISIS  00h11m54s 15
                                fe80::10:1ff:fe01:1-"int-P-3-PE-1"
                                20
2001:db8:110::2/128             Remote ISIS  00h00m15s 15
                                fe80::10:1ff:fe01:1-"int-P-3-PE-1"
                                10
2001:db8:110::3/128             Remote ISIS  00h00m15s 15
                                fe80::10:1ff:fe01:1-"int-P-3-PE-1"
                                10
2001:db8:110::4/128             Remote ISIS  00h00m15s 15
                                fe80::10:1ff:fe01:1-"int-P-3-PE-1"
                                10
-----
No. of Routes: 6
Flags: n = Number of times nexthop is repeated
      B = BGP backup route available
      L = LFA nexthop available
      S = Sticky ECMP requested
=====
```

The export policy is removed for IS-IS, as follows:

```
# on PE-1:
configure
router Base
isis 0
no export
```

## Export IPv6 VPN-leak routes to OSPFv3

The export policy for IPv6 routes of protocol VPN-leak is applied for OSPFv3, as follows:

```
# on PE-1:
configure
router Base
ospf3
export "export-vpn-leak"
```

Routes can only be exported to OSPFv3 if the router is configured as ASBR, as follows:

```
# on PE-1:
configure
router Base
ospf3
asbr
```

The IPv6 VPN-leak routes from PE-1 are now advertised by OSPFv3 to PE-2 and P-3. The preference for remote OSPFv3 routes is by default 150. The routing table on P-3 contains the following IPv6 OSPFv3 routes:

```
*A:P-3# show router route-table ipv6 protocol ospf3
=====
IPv6 Route Table (Router: Base)
=====
Dest Prefix[Flags]
Next Hop[Interface Name]          Type   Proto   Age           Pref
Metric
-----
2001:db8::1/128
  fe80::10:1ff:fe01:1-"int-P-3-PE-1" Remote OSPF3   00h00m48s    10
  10
2001:db8::2/128
  fe80::14:1ff:fe01:2-"int-P-3-PE-2" Remote OSPF3   00h00m48s    10
  10
2001:db8:12::/126
  fe80::10:1ff:fe01:1-"int-P-3-PE-1" Remote OSPF3   00h00m48s    10
  20
2001:db8:110::2/128
  fe80::10:1ff:fe01:1-"int-P-3-PE-1" Remote OSPF3   00h00m22s    150
  1
2001:db8:110::3/128
  fe80::10:1ff:fe01:1-"int-P-3-PE-1" Remote OSPF3   00h00m22s    150
  1
2001:db8:110::4/128
  fe80::10:1ff:fe01:1-"int-P-3-PE-1" Remote OSPF3   00h00m22s    150
  1
-----
No. of Routes: 6
Flags: n = Number of times nexthop is repeated
      B = BGP backup route available
      L = LFA nexthop available
      S = Sticky ECMP requested
=====
```

The export policy is removed for OSPFv3, as follows:

```
# on PE-1:
configure
  router Base
    ospf3
      no export
```

## Export IPv6 VPN-leak routes to BGP

The export policy for IPv6 routes of protocol VPN-leak is applied for BGP, as follows:

```
# on PE-1:
configure
  router Base
    bgp
      export "export-vpn-leak"
```

The three IPv6 VPN-leak routes from PE-1 are now advertised by BGP to PE-2 and P-3. The routing table on P-3 contains the following IPv6 BGP routes:

```
*A:P-3# show router route-table ipv6 protocol bgp
=====
IPv6 Route Table (Router: Base)
=====
```

Dest Prefix[Flags] Next Hop[Interface Name]	Type	Proto	Age Metric	Pref
2001:db8:110::2/128 fe80::10:1ff:fe01:1-"int-P-3-PE-1"	Remote	BGP	00h00m15s 10	170
2001:db8:110::3/128 fe80::10:1ff:fe01:1-"int-P-3-PE-1"	Remote	BGP	00h00m15s 10	170
2001:db8:110::4/128 fe80::10:1ff:fe01:1-"int-P-3-PE-1"	Remote	BGP	00h00m15s 10	170

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

The export policy is removed for BGP, as follows:

```
# on PE-1:
configure
  router Base
    bgp
      no export
```

In this example, BGP leaked IPv6 routes are advertised by BGP. For scenarios with only BGP routes, a dedicated BGP route leaking mechanism that preserves all attributes is preferred, as described in chapter *BGP Route Leaking*. However, with the same configuration as in this chapter, it is possible to leak non-BGP routes and advertise them using BGP.

## Conclusion

Routes learned in a VPRN can be leaked to the base router and advertised using routing protocols. The mechanism described in this chapter is protocol-independent: all kinds of routes can be leaked from a VRF to the GRT: local, static, IS-IS, OSPF, BGP routes, and so on. In some cases, it might be useful to leak the routes from a VPRN to the entire network using the routing protocol, in order to access the resources defined inside the VRF. Routes that are leaked from VPRNs to the GRT must be unique in the network where they will be advertised.

For BGP routes, the protocol-independent route leaking mechanism described here does not preserve the attributes, unlike the dedicated BGP route leaking feature.

# Weighted ECMP for VPRN over RSVP-TE or SR-TE LSPs

This chapter provides information about Weighted Equal Cost Multipath (ECMP) for Virtual Private Routed Network (VPRN) over Resource Reservation Protocol with Traffic Engineering (RSVP-TE) or Segment Routing with Traffic Engineering (SR-TE) Label Switched Paths (LSPs).

Topics in this chapter include:

- [Applicability](#)
- [Overview](#)
- [Configuration](#)
- [Conclusion](#)

## Applicability

The information and configuration in this chapter are based on SR OS Release 23.3.R2. Weighted load balancing over Multi Protocol Label Switching (MPLS) LSPs - as described in chapter *BGP Weighted ECMP* in the *7450 ESS, 7750 SR, and 7950 XRS Unicast Routing Protocols Advanced Configuration Guide for Classic CLI* - is supported in SR OS Release 13.0.R1, and later. Weighted load balancing for VPRN with auto-bind-tunnel over RSVP-TE LSPs is supported in SR OS Release 15.0.R2, and later. Weighted load balancing for VPRN with auto-bind-tunnel over SR-TE LSPs is supported in SR OS Release 15.0.R4, and later.

## Overview

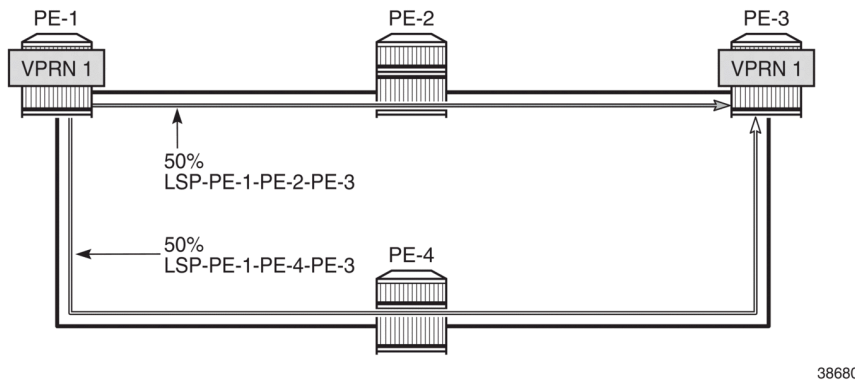
### Equal Load Balancing

In this chapter, ECMP refers to spraying traffic flows over multiple RSVP-TE or SR-TE LSPs within an ECMP set. ECMP spraying consists of hashing the relevant fields in the packet header and selecting the tunnel next-hop based on the modulo operation of the output of the hash and the number of LSPs present in the ECMP set. The maximum number of LSPs in the ECMP set is defined by the **ecmp** command.

Only LSPs with the same lowest LSP metric can be part of the ECMP set. If the number of such LSPs exceeds the maximum number of LSPs allowed in the ECMP set as defined by the **ecmp** command, the LSPs with the lowest tunnel IDs are selected first. By default, all LSPs in the ECMP set have the same weight, and traffic flows are spread evenly over all LSPs in the ECMP set, regardless of the bandwidth of the active path in the LSPs. By default, ECMP is enabled and set to 1.

**Figure 93: Regular ECMP in AS 64496** shows that PE-1 sprays the traffic flows equally over two LSPs between PE-1 and PE-3. If three or more LSPs with the same lowest LSP metric were available from PE-1 to PE-3, only two of those would be used, because an ECMP value of 2 allows the traffic to be sprayed over two LSPs.

Figure 93: Regular ECMP in AS 64496

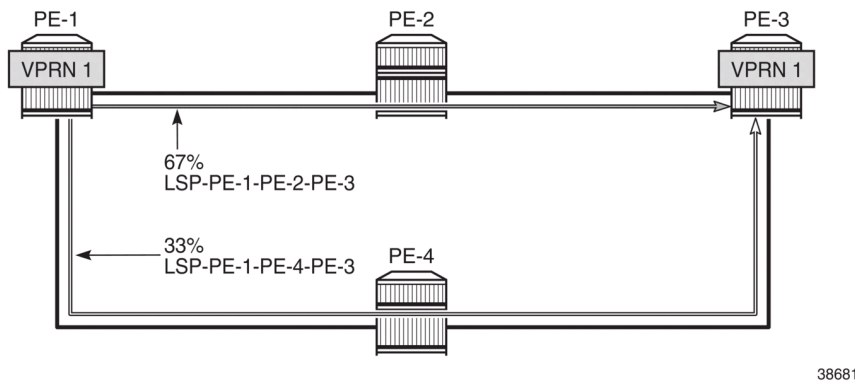


38680

## Unequal Load Balancing

Weighted ECMP sprays traffic flows over MPLS LSPs proportionally to the **load-balancing-weight** `<weight>` value configured on each MPLS LSP in the ECMP set. [Figure 94: Weighted ECMP in AS 64496](#) shows that PE-1 forwards two thirds of the traffic flows on LSP-PE-1-PE-2-PE-3 with weight 2 and one third on LSP-PE-1-PE-4-PE-3 with weight 1. Each of the links can be link aggregation group (LAG) ports. For instance, when LSP-PE-1-PE-2-PE-3 uses LAG ports, 67% of the traffic is sprayed evenly over all ports belonging to the LAG.

Figure 94: Weighted ECMP in AS 64496



38681

The LSP load balancing weight can be configured in an LSP template or on an LSP. By default, the load balancing weight equals zero, in which case regular ECMP applies.

The following command is used to configure the weight in an LSP template:

```
*A:PE-1# configurerouter Base mpls lsp-template "LSPtemplate1" load-balancing-weight ?
- no load-balancing-weight
- load-balancing-weight <weight>

<weight>          : [0..4294967295] Default - 0
```

The following command is used to configure the weight on an LSP (for example on LSP "LSP-PE-1-PE-2-PE-3"):

```
*A:PE-1# configurerouter Base mpls lsp "LSP-PE-1-PE-2-PE-3" load-balancing-weight ?
- load-balancing-weight <weight>
- no load-balancing-weight

<weight>          : [0..4294967295] Default - 0
```

The LSP load balancing weight on LSP-PE-1-PE-2-PE-3 is configured with a value of 2, as follows:

```
configure
router Base
 mpls
  path "path-PE-1-PE-2-PE-3"
    hop 10 192.168.12.2 strict
    hop 20 192.168.23.2 strict
    no shutdown
  exit
  lsp "LSP-PE-1-PE-2-PE-3"
    to 192.0.2.3
    path-computation-method local-cspf
    metric 100
    load-balancing-weight 2
    primary "path-PE-1-PE-2-PE-3"
  exit
  no shutdown
exit
```

Weighted ECMP is enabled in the **vprn 1 bgp-ipvpn mpls auto-bind-tunnel** context as follows:

```
configure
service
 vprn 1 name "1" customer 1 create
  bgp-ipvpn
  mpls
    auto-bind-tunnel
    ecmp 2
    weighted-ecmp
  exit
```

Weighted load balancing within a **vprn** context can be performed only when the next-hops are associated with the same neighbor and all LSPs in the ECMP set are configured with non-zero load balancing weights. If one or more LSPs in the ECMP set toward a specific next-hop do not have a load balancing weight configured, regular ECMP spraying is used. The weighted ECMP support for ECMP routes applies to both IPv4 and IPv6.

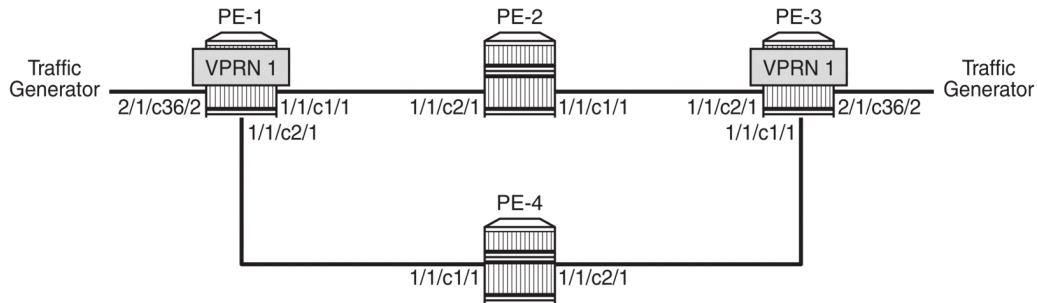
Additionally, it is possible to enable ECMP in the **vprn** context, with: **configure service vprn 1 ecmp <max-ecmp-routes>**, to control load balancing to a different next-hop. The **weighted-ecmp** option in the **VPN 1 bgp-ipvpn mpls auto-bind-tunnel** context controls load balancing to the same next-hop.

## Configuration

[Figure 95: Example Topology](#) shows the example topology with four PEs. VPRN 1 is configured on PE-1 and PE-3. A traffic generator is connected to VPRN 1 SAP 2/1/c36/2 on PE-1 and VPRN 1 SAP 2/1/c36/2

on PE-3. The traffic generator generates multiple traffic flows with random IP addresses and TCP/UDP port numbers. As a result, these flows are sprayed over different MPLS LSPs between PE-1 and PE-3.

Figure 95: Example Topology



38676

The initial configuration on the PEs includes the following:

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

The initial configuration on PE-1 is as follows:

```
configure
router Base
  interface "int-PE-1-PE-2"
    address 192.168.12.1/30
    port 1/1/c1/1
  exit
  interface "int-PE-1-PE-4"
    address 192.168.14.1/30
    port 1/1/c2/1
  exit
  interface "system"
    address 192.0.2.1/32
  exit
  isis 0
    area-id 49.0001
    traffic-engineering
    interface "system"
    exit
    interface "int-PE-1-PE-2"
      interface-type point-to-point
    exit
    interface "int-PE-1-PE-4"
      interface-type point-to-point
    exit
    no shutdown
  exit
  mpls
    interface "int-PE-1-PE-2"
    exit
    interface "int-PE-1-PE-4"
    exit
    no shutdown
```



```

exit
 rsvp
  no shutdown
exit

```

The initial configuration on the other PEs is similar.

With the preceding configuration, MPLS and RSVP are enabled on all interfaces, including the system interface, which is added automatically.

In the next sections, the following use cases are described:

- [Weighted ECMP for VPRN with Auto-bind-tunnel RSVP-TE](#)
- [Weighted ECMP for an SDP used as a spoke SDP in a VPRN](#)
- [Weighted ECMP for VPRN with Auto-bind-tunnel SR-TE](#)

## Weighted ECMP for VPRN with Auto-bind-tunnel RSVP-TE

On PE-1, the following paths and LSPs are configured. LSP-PE-1-PE-2-PE-3 is configured with a load balancing weight of 2; LSP-PE-1-PE-4-PE-3 is configured with a load balancing weight of 1.

```

configure
  router Base
    mpls
      path "path-PE-1-PE-2-PE-3"
        hop 10 192.168.12.2 strict
        hop 20 192.168.23.2 strict
        no shutdown
      exit
      path "path-PE-1-PE-4-PE-3"
        hop 10 192.168.14.2 strict
        hop 20 192.168.34.1 strict
        no shutdown
      exit
      lsp "LSP-PE-1-PE-2-PE-3"
        to 192.0.2.3
        path-computation-method local-cspf
        metric 100
        load-balancing-weight 2
        primary "path-PE-1-PE-2-PE-3"
        exit
        no shutdown
      exit
      lsp "LSP-PE-1-PE-4-PE-3"
        to 192.0.2.3
        path-computation-method local-cspf
        metric 100
        load-balancing-weight 1
        primary "path-PE-1-PE-4-PE-3"
        exit
        no shutdown
      exit
    exit
  exit

```

On PE-1, VPRN 1 is configured as follows. ECMP and weighted ECMP can be configured in the **vprn** context, for example, **configure service vprn 1 ecmp 2** and **configure service vprn 1 weighted-ecmp** but it is not required when the next-hop for the MPLS LSPs is the same. In this example, ECMP and weighted ECMP are only configured in the **vprn 1 bgp-ipvprn mpls auto-bind-tunnel** context. The

resolution filter only allows RSVP-TE tunnels, no other MPLS LSPs, such as LDP, BGP, or segment routing (SR) tunnels.

```
configure
  service
    vprn 1 name "1" customer 1 create
      description "CE-1"
      bgp-ipvpn
      mpls
        auto-bind-tunnel
        resolution-filter
        rsvp
        exit
        resolution filter
        ecmp 2
        weighted-ecmp
        exit
        route-distinguisher 64496:1
        vrf-target target:64496:1
        no shutdown
      exit
    exit
  interface "loopback1" create
    address 172.16.0.1/32
    ipv6
      address 2001:db8::1/128
    exit
  loopback
  exit
  interface "int-CE-1-STC" create
    address 192.168.11.1/24
    ipv6
      address 2001:db8::11:1/120
    exit
    sap 2/1/c36/2 create
    exit
  exit
  no shutdown
exit
```

The service configuration on PE-3 is similar.

VPRN 1 is dual stacked. Weighted ECMP applies to both IPv4 and IPv6 traffic streams. BGP is configured for the VPN-IPv4 and VPN-IPv6 address family to exchange the routes used in VPRN 1 between PE-1 and PE-3. The BGP configuration on PE-1 is as follows:

```
configure
  router Base
    autonomous-system 64496
    bgp
      group "iBGP"
        neighbor 192.0.2.3
          family vpn-ipv4 vpn-ipv6
          export "export-vpn-ipv4" "export-vpn-ipv6"
          peer-as 64496
        exit
    exit
```

The BGP configuration on PE-3 is similar.

The export policies on PE-1 are defined as follows:

```
configure
```

```

router Base
  policy-options
  begin
    prefix-list "vpn-ipv4"
      prefix 172.16.0.0/16 longer
      prefix 192.168.11.0/24 exact
    exit
    prefix-list "vpn-ipv6"
      prefix 2001:db8::/120 longer
      prefix 2001:db8::11:0/120 exact
    exit
    policy-statement "export-vpn-ipv4"
      entry 10
        from
          prefix-list "vpn-ipv4"
        exit
        action accept
        exit
      exit
    exit
    policy-statement "export-vpn-ipv6"
      entry 10
        from
          prefix-list "vpn-ipv6"
        exit
        action accept
        exit
      exit
    exit
  exit
  commit
  
```

The export policies on PE-3 are similar.

With ECMP enabled for MPLS LSPs with the same next-hop and two RSVP-TE LSPs available with equal metric, the route table of VPRN 1 on PE-1 shows two routes for each prefix with the same next-hop 192.0.2.3: one via RSVP LSP 1 and the other via RSVP LSP 2, as follows:

```
*A:PE-1# show router 1 route-table
```

```
=====
Route Table (Service: 1)
=====
```

Dest Prefix[Flags] Next Hop[Interface Name]	Type	Proto	Age	Metric	Pref
172.16.0.1/32 loopback1	Local	Local	00h01m16s	0	0
172.16.0.3/32 [2] 192.0.2.3 (tunneled:RSVP:1)	Remote	BGP VPN	00h00m29s	100	170
172.16.0.3/32 [2] 192.0.2.3 (tunneled:RSVP:2)	Remote	BGP VPN	00h00m29s	100	170
192.168.11.0/24 int-CE-1-STC	Local	Local	00h01m16s	0	0
192.168.33.0/24 [2] 192.0.2.3 (tunneled:RSVP:1)	Remote	BGP VPN	00h00m29s	100	170
192.168.33.0/24 [2] 192.0.2.3 (tunneled:RSVP:2)	Remote	BGP VPN	00h00m29s	100	170

```
-----
No. of Routes: 6
```

```
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
```

The flag [2] indicates that next-hop 192.0.2.3 occurs twice for the prefix 172.16.0.3/32; next-hop 192.0.2.3 also occurs twice for the prefix 192.168.33.0/24.

The following IPv6 route table is similar, with next-hop 192.0.2.3 occurring twice for prefix 2001:db8::3/128 and twice for prefix 2001:db8::33:0/120.

```
*A:PE-1# show router 1 route-table ipv6

=====
IPv6 Route Table (Service: 1)
=====
Dest Prefix[Flags]                               Type  Proto  Age           Pref
  Next Hop[Interface Name]                       Metric
-----
2001:db8::1/128                                  Local  Local   00h01m15s    0
  loopback1
2001:db8::3/128 [2]                               Remote BGP VPN 00h00m29s    170
  192.0.2.3 (tunneled:RSVP:1)                    100
2001:db8::3/128 [2]                               Remote BGP VPN 00h00m29s    170
  192.0.2.3 (tunneled:RSVP:2)                    100
2001:db8::11:0/120                               Local  Local   00h01m15s    0
  int-CE-1-STC
2001:db8::33:0/120 [2]                           Remote BGP VPN 00h00m29s    170
  192.0.2.3 (tunneled:RSVP:1)                    100
2001:db8::33:0/120 [2]                           Remote BGP VPN 00h00m29s    170
  192.0.2.3 (tunneled:RSVP:2)                    100
-----
No. of Routes: 6
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
=====
```

The following tunnel table output on PE-1 shows that RSVP-TE LSP 1 goes via PE-2 (next-hop 192.168.12.2) and RSVP-TE LSP 2 via PE-4 (next-hop 192.168.14.2):

```
*A:PE-1# show router tunnel-table

=====
IPv4 Tunnel Table (Router: Base)
=====
Destination      Owner    Encap TunnelId  Pref  Nexthop      Metric
  Color
-----
192.0.2.3/32     rsvp    MPLS  1           7     192.168.12.2  100
192.0.2.3/32     rsvp    MPLS  2           7     192.168.14.2  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
=====
```

To verify the weighted load balancing between the two RSVP-TE LSPs, the traffic generator sends multiple IPv4 and IPv6 traffic flows with random IP addresses and TCP/UDP port numbers via PE-1 to PE-3. The traffic enters PE-1 through port 2/1/c36/2. When LSP-PE-1-PE-2-PE-3 is configured with weight 2 and

LSP-PE-1-PE-4-PE-3 with weight 1, PE-1 forwards two thirds of the traffic via port 1/1/c1/1 toward PE-2 and one third of the traffic via port 1/1/c2/1 toward PE-3, as follows:

```
*A:PE-1# monitor port 1/1/c1/1 1/1/c2/1 2/1/c36/2 rate interval 3 repeat 3

=====
Monitor statistics for Ports
=====
-----
Input                               Output
-----
---snip---
-----
At time t = 6 sec (Mode: Rate)
-----
Port 1/1/c1/1
-----
Octets                               139                               431549
Packets                              2                               419
Errors                               0                               0
Bits                                 1112                            3452392
Utilization (% of port capacity)    ~0.00                            0.03

Port 1/1/c2/1
-----
Octets                               46                               180781
Packets                              0                               177
---snip---

Port 2/1/c36/2
-----
Octets                               608256                           0
Packets                              594                            0
---snip---
=====
```

This can also be verified as follows:

```
*A:PE-1# show port 1/1/c1/1 statistics

=====
Port Statistics on Slot 1
=====
Port                               Ingress Packets                    Ingress Octets
Id                                 Egress Packets                     Egress Octets
-----
1/1/c1/1                           42                                  4121
                                   10900                             11209443
=====

*A:PE-1# show port 1/1/c2/1 statistics

=====
Port Statistics on Slot 1
=====
Port                               Ingress Packets                    Ingress Octets
Id                                 Egress Packets                     Egress Octets
-----
1/1/c2/1                           36                                  3557
                                   4549                             4661813
=====

*A:PE-1# show port 2/1/c36/2 statistics
```

```

=====
Port Statistics on Slot 2
=====
Port          Ingress Packets      Ingress Octets
Id            Egress Packets      Egress Octets
-----
2/1/c36/2    15372                15740928
              0                    0
=====
    
```

## Restrictions

All RSVP-TE LSPs in the ECMP set must have a load balancing weight configured. When at least one RSVP-TE LSP in the ECMP set is configured without weight, regular ECMP is applied.

## RSVP-TE LSP without Weight in ECMP Set

If one of the RSVP-TE LSPs in the ECMP set does not have a load balancing weight configured, the traffic flows are sprayed equally between all RSVP-TE LSPs, regardless of the configured weight of the other RSVP-TE LSPs in the ECMP set.

On PE-1, LSP-PE-1-PE-2-PE-3 is configured without a load balancing weight, as follows:

```

configure
router Base
 mpls
  lsp "LSP-PE-1-PE-2-PE-3"
   no load-balancing-weight
    
```

LSP-PE-1-PE-4-PE-3 is still configured with a load balancing weight of 1, but it is impossible to calculate its relative weight, because the sum of the weight values is not defined. Therefore, PE-1 reverts to regular ECMP for the load balancing between the two RSVP-TE LSPs to PE-3. When the traffic generator sends multiple traffic flows via PE-1 to PE-3, the load is spread equally over both RSVP-TE LSPs, as shown in the following monitor output. Port 1/1/c1/1 is used for traffic sent via LSP-PE-1-PE-2-PE-3 and port 1/1/c2/1 for traffic sent via LSP-PE-1-PE-4-PE-3.

```

*A:PE-1# monitor port 1/1/c1/1 1/1/c2/1 2/1/c36/2 rate interval 3 repeat 3
=====
Monitor statistics for Ports
=====
                                Input          Output
-----
---snip---
-----
At time t = 6 sec (Mode: Rate)
-----
Port 1/1/c1/1
-----
Octets                21                304157
Packets                0                  295
Errors                 0                   0
Bits                  168                2433256
Utilization (% of port capacity) ~0.00              0.02

Port 1/1/c2/1
-----
    
```

```

Octets                21                305837
Packets               0                  297
---snip---

Port 2/1/c36/2
-----
Octets                605867                0
Packets              592                  0
---snip---
=====

```

The configuration is restored as follows:

```

configure
  router Base
    mpls
      lsp "LSP-PE-1-PE-2-PE-3"
        load-balancing-weight 2

```

## Weighted ECMP for an SDP used as a spoke SDP in a VPRN

The following LSPs are configured on PE-1. The LSP load balancing weight values are 4 and 1 and the metric is 101 for both LSPs.

```

configure
  router Base
    mpls
      lsp "LSP-PE-1-PE-2-PE-3-spoke"
        to 192.0.2.3
        path-computation-method local-cspf
        metric 101
        load-balancing-weight 4
        primary "path-PE-1-PE-2-PE-3"
        exit
        no shutdown
      exit
      lsp "LSP-PE-1-PE-4-PE-3-spoke"
        to 192.0.2.3
        path-computation-method local-cspf
        metric 101
        load-balancing-weight 1
        primary "path-PE-1-PE-4-PE-3"
        exit
        no shutdown
      exit
    exit
  no shutdown
exit
  rsvp
  no shutdown
exit

```

Similar LSPs are configured on PE-3.

On PE-1, an SDP is configured, as follows:

```

configure
  service
    sdp 13 mpls create
      far-end 192.0.2.3

```

```

lsp "LSP-PE-1-PE-2-PE-3-spoke"
lsp "LSP-PE-1-PE-4-PE-3-spoke"
no shutdown
exit

```

A similar SDP is configured on PE-3.

These SDPs are configured as spoke SDPs in a VPRN, as follows:

```

configure
service
vprn 1
spoke-sdp 13 create
exit

configure
router Base
ldp
no shutdown
exit

```

On PE-1, weighted ECMP is enabled on an SDP, as follows:

```

configure
service
sdp 13
weighted-ecmp
exit

```

The ECMP configuration on PE-3 is similar.

With ECMP enabled for MPLS LSPs with the same next-hop and two RSVP-TE LSPs available with equal metric, the route table of VPRN 1 on PE-1 shows one route for each prefix via the SDP tunnel, as follows:

```

*A:PE-1# show router 1 route-table

=====
Route Table (Service: 1)
=====
Dest Prefix[Flags]
Next Hop[Interface Name]
Type Proto Age Metric Pref
-----
172.16.0.1/32
loopback1 Local Local 00h21m51s 0
172.16.0.3/32
192.0.2.3 (tunneled) Remote BGP VPN 00h01m31s 170
192.168.11.0/24
int-CE-1-STC Local Local 00h21m51s 0
192.168.33.0/24
192.0.2.3 (tunneled) Remote BGP VPN 00h01m31s 170
-----
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
=====

```



The following tunnel table output on PE-1 shows that the preferred route to PE-3 is via the SDP tunnel. It uses RSVP-TE LSP 3 that goes via PE-2 (next-hop 192.168.12.2) and RSVP-TE LSP 4 that goes via PE-4 (next-hop 192.168.14.2). :

```
*A:PE-1# show router tunnel-table

=====
IPv4 Tunnel Table (Router: Base)
=====
Destination          Owner    Encap TunnelId  Pref  Nexthop      Metric
  Color
-----
192.0.2.3/32         sdp      MPLS   13       5    192.0.2.3      0
---snip---
192.0.2.3/32         rsvp     MPLS    3       7    192.168.12.2  101
192.0.2.3/32         rsvp     MPLS    4       7    192.168.14.2  101
-----
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
=====
```

To verify the weighted load balancing between the two RSVP-TE LSPs, the traffic generator sends multiple IPv4 traffic flows with random IP addresses and TCP/UDP port numbers via PE-1 to PE-3. The traffic enters PE-1 through port 2/1/c36/2. When LSP-PE-1-PE-2-PE-3-spoke is configured with weight 4 and LSP-PE-1-PE-4-PE-3-spoke with weight 1, PE-1 forwards four fifths of the traffic via port 1/1/c1/1 toward PE-2 and one fifth of the traffic via port 1/1/c2/1 toward PE-3, as follows:

```
*A:PE-1# monitor port 1/1/c1/1 1/1/c2/1 2/1/c36/2 rate interval 3 repeat 3

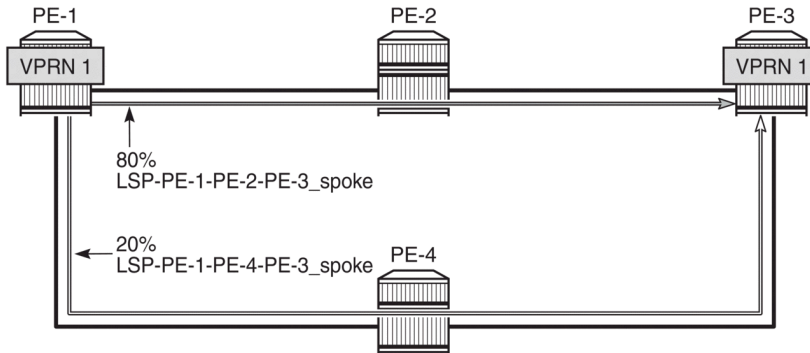
=====
Monitor statistics for Ports
=====
                                     Input          Output
-----
---snip---
At time t = 6 sec (Mode: Rate)
-----
Port 1/1/c1/1
-----
Octets                               74              489052
Packets                              1               476
Errors                               0                0
Bits                                592             3912416
Utilization (% of port capacity)    ~0.00            0.03

Port 1/1/c2/1
-----
Octets                               61              118053
Packets                              0               115
---snip---

Port 2/1/c36/2
-----
Octets                               602112           0
Packets                              588              0
---snip---
=====
```

Figure 96: Weighted ECMP over RSVP LSPs used in a spoke SDP shows how the traffic flows are sprayed over the two RSVP LSPs:

Figure 96: Weighted ECMP over RSVP LSPs used in a spoke SDP



38679

## Weighted ECMP for VPRN with Auto-bind-tunnel SR-TE

The following configuration is added to enable SR-ISIS on PE-1.

```
configure
  router Base
    mpls-labels
      sr-labels start 20000 end 20099
    exit
    isis 0
      advertise-router-capability as
      interface "system"
        ipv4-node-sid label 20001
      exit
      segment-routing
        prefix-sid-range start-label 20000 max-index 99
        no shutdown
      exit
    exit
```

The configuration on the other PEs is identical, except for the **ipv4-node-sid label** value.

The following SR-TE LSPs are configured on PE-1. For more information about SR-TE LSPs, see the *Segment Routing Traffic Engineered Tunnels* chapter. The load balancing weight values are 75 and 25. The values 3 and 1, which have the same ratio, can be used instead.

```
configure
  router Base
    mpls
      lsp "LSP-PE-1-PE-2-PE-3_SR-TE" sr-te
        to 192.0.2.3
        load-balancing-weight 75
        primary "path-PE-1-PE-2-PE-3"
        exit
        no shutdown
      exit
      lsp "LSP-PE-1-PE-4-PE-3_SR-TE" sr-te
```

```

to 192.0.2.3
  load-balancing-weight 25
  primary "path-PE-1-PE-4-PE-3"
exit
no shutdown
exit

```

The configuration on PE-3 is similar.

The following tunnel table on PE-1 shows two SR-TE tunnels with equal metrics: SR-TE tunnel 655362 has PE-2 as next-hop (192.168.12.2) and SR-TE tunnel 655363 has PE-4 as next-hop (192.168.14.2).

```

*A:PE-1# show router tunnel-table protocol sr-te
=====
IPv4 Tunnel Table (Router: Base)
=====
Destination      Owner      Encap TunnelId  Pref  Nexthop      Metric
  Color
-----
192.0.2.3/32     sr-te     MPLS  655362    8    192.168.12.2  16777215
192.0.2.3/32     sr-te     MPLS  655363    8    192.168.14.2  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 resolution filter for VPRN 1 is configured on PE-1 and PE-3 to only allow SR-TE tunnels. ECMP and weighted ECMP are enabled in the **vprn 1 bgp-ipvpn mpls auto-bind-tunnel** context.

```

configure
  service
    vprn 1 name "1" customer 1 create
    description "CE-1"
    bgp-ipvpn
      mpls
        auto-bind-tunnel
          resolution-filter
            sr-te
          exit
          resolution filter
            ecmp 2
            weighted-ecmp
          exit
          route-distinguisher 64496:1
          vrf-target target:64496:1
          no shutdown
        exit
      exit
    interface "loopback1" create
      address 172.16.0.1/32
      ipv6
        address 2001:db8::1/128
      exit
      loopback
    exit
    interface "int-CE-1-STC" create
      address 192.168.11.1/24
      ipv6
        address 2001:db8::11:1/120

```

```

        exit
        sap 2/1/c36/2 create
        exit
    exit
    no shutdown
exit
    
```

The following route table for VPRN 1 on PE-1 shows two entries for each remote prefix with the same next-hop 192.0.2.3 and a different SR-TE LSP: SR-TE tunnel 655362 and SR-TE tunnel 655363.

```

*A:PE-1# show router 1 route-table

=====
Route Table (Service: 1)
=====
Dest Prefix[Flags]          Type  Proto  Age      Pref
  Next Hop[Interface Name]      Metric
-----
172.16.0.1/32              Local  Local   00h39m10s  0
  loopback1                  0
172.16.0.3/32 [2]          Remote BGP VPN 00h01m15s 170
  192.0.2.3 (tunneled:SR-TE:655362) 16777215
172.16.0.3/32 [2]          Remote BGP VPN 00h01m15s 170
  192.0.2.3 (tunneled:SR-TE:655363) 16777215
192.168.11.0/24            Local  Local   00h39m10s  0
  int-CE-1-STC                0
192.168.33.0/24 [2]        Remote BGP VPN 00h01m15s 170
  192.0.2.3 (tunneled:SR-TE:655362) 16777215
192.168.33.0/24 [2]        Remote BGP VPN 00h01m15s 170
  192.0.2.3 (tunneled:SR-TE:655363) 16777215
-----
No. of Routes: 6
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
=====
    
```

The following IPv6 route table for VPRN 1 is similar:

```

*A:PE-1# show router 1 route-table ipv6

=====
IPv6 Route Table (Service: 1)
=====
Dest Prefix[Flags]          Type  Proto  Age      Pref
  Next Hop[Interface Name]      Metric
-----
2001:db8::1/128            Local  Local   00h39m09s  0
  loopback1                  0
2001:db8::3/128 [2]          Remote BGP VPN 00h01m15s 170
  192.0.2.3 (tunneled:SR-TE:655362) 16777215
2001:db8::3/128 [2]          Remote BGP VPN 00h01m15s 170
  192.0.2.3 (tunneled:SR-TE:655363) 16777215
2001:db8::11:0/120         Local  Local   00h39m09s  0
  int-CE-1-STC                0
2001:db8::33:0/120 [2]      Remote BGP VPN 00h01m15s 170
  192.0.2.3 (tunneled:SR-TE:655362) 16777215
2001:db8::33:0/120 [2]      Remote BGP VPN 00h01m15s 170
  192.0.2.3 (tunneled:SR-TE:655363) 16777215
-----
No. of Routes: 6
    
```

```
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
=====
```

When multiple IPv4 and IPv6 traffic flows are sent from PE-1 to PE-3, the load balancing is weighted: 75% is sent via port 1/1/c1/1 toward PE-2 (LSP-PE-1-PE-2-PE-3\_SR-TE) and 25% is sent via port 1/1/c2/1 toward PE-4 (LSP-PE-1-PE-4-PE-3\_SR-TE), as follows:

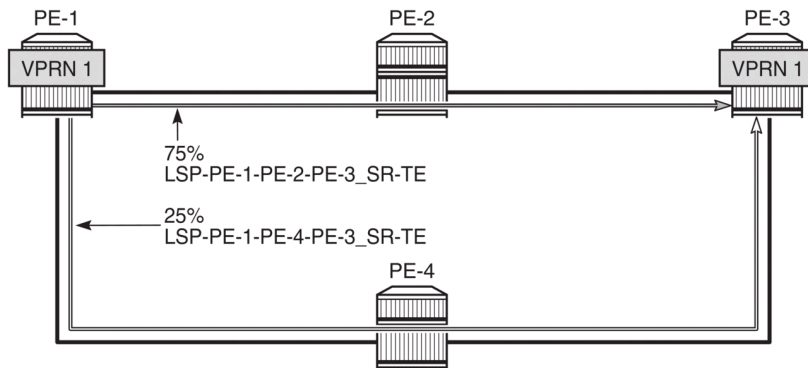
```
*A:PE-1# monitor port 1/1/c1/1 1/1/c2/1 2/1/c36/2 rate interval 3 repeat 3
=====
Monitor statistics for Ports
=====
                                     Input           Output
-----
---snip---
-----
At time t = 6 sec (Mode: Rate)
-----
Port 1/1/c1/1
-----
Octets                98                453453
Packets                1                440
Errors                 0                 0
Bits                  784              3627624
Utilization (% of port capacity)  ~0.00             0.03

Port 1/1/c2/1
-----
Octets                43                154231
Packets                1                150
---snip---

Port 2/1/c36/2
-----
Octets               595968              0
Packets              582                0
---snip---
```

Figure 97: Weighted ECMP over SR-TE LSPs in AS 64496 shows how the traffic flows are sprayed over the two SR-TE LSPs:

Figure 97: Weighted ECMP over SR-TE LSPs in AS 64496



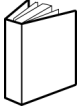
38682

## Conclusion

Operators can control how traffic in a VPRN is load balanced unequally over multiple transport tunnels by defining a load balancing weight factor on each LSP and enabling weighted ECMP in the VPRN. In this chapter, weighted ECMP for VPRN over transport LSPs is enabled for RSVP-TE tunnels and for SR-TE tunnels.



# Customer document and product support



## **Customer documentation**

[Customer documentation welcome page](#)



## **Technical support**

[Product support portal](#)



## **Documentation feedback**

[Customer documentation feedback](#)